ProHits Analyst

User manual - demo

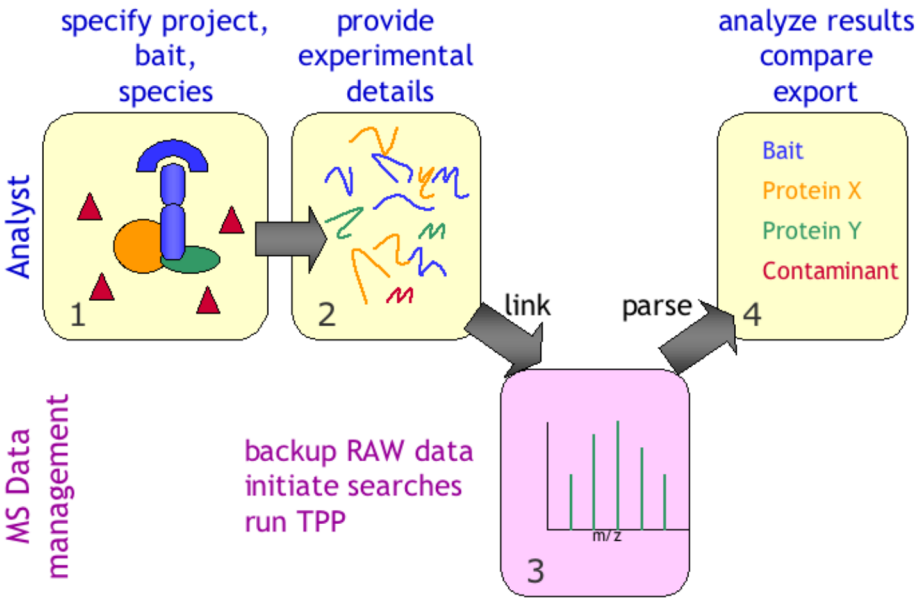
Version demo1, prepared on Jan 15, 2010, by Anne-Claude Gingras, with the help of Frank Liu, JP Zhang, Brian Raught, Brett Larsen, Wade Dunham, Marilyn Goudreault and Karen Colwill.

Contents

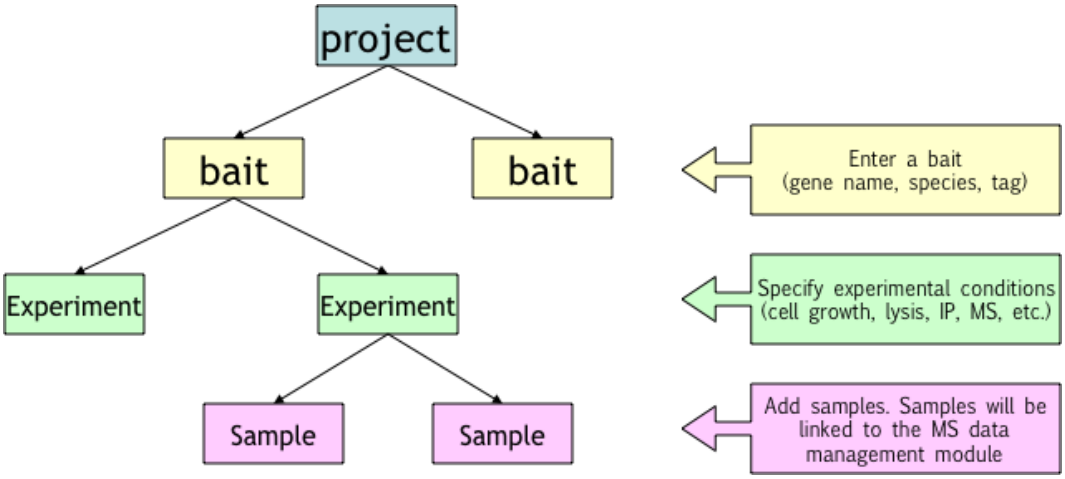
Overview 2
Access to projects 3
Analyst main page 4
Description of the navigator bar options: 4
Creating samples and viewing individual reports 5
Adding a "Gel-Free" sample 5
Creating a bait 5
Creating an experiment 8
User-defined free-text protocols 8
Controlled vocabulary 8
Additional annotation 8
Creating a new sample 11
Linking raw files from the Data Management module 12
Linking raw files directly from the Analyst module (alternative) 12
Navigating through the results 15
Search results 15
Sorting options 16
Links details 16
Option details 16
Using filters 17
Comparing your data to literature interactions 19
View and navigate hits from the TransProteomics Pipeline 20
Viewing results using Cytoscape 21
Export Sample report 21
Using the Notes option 22
Creating gel-based samples 23
Adding a "Gel-based" sample 23
The Comparison tool 26
Merging files prior to Comparison 27
Comparison page 29
Using Cytoscape directly from ProHits comparison 31
Other export options 32
Comparing larger numbers of baits 33
Automatically adding baits for comparison from the baits or sample report list pages 34
Simple Search (Gene name) 35
Other keywords that can be searched: 37
Manage Protocols and Lists 41
Text-based protocols 42
Experimental Editor 44
Background Lists 46
Group Lists 49
Export version 51
Epitope Tag Lists 52

Overview

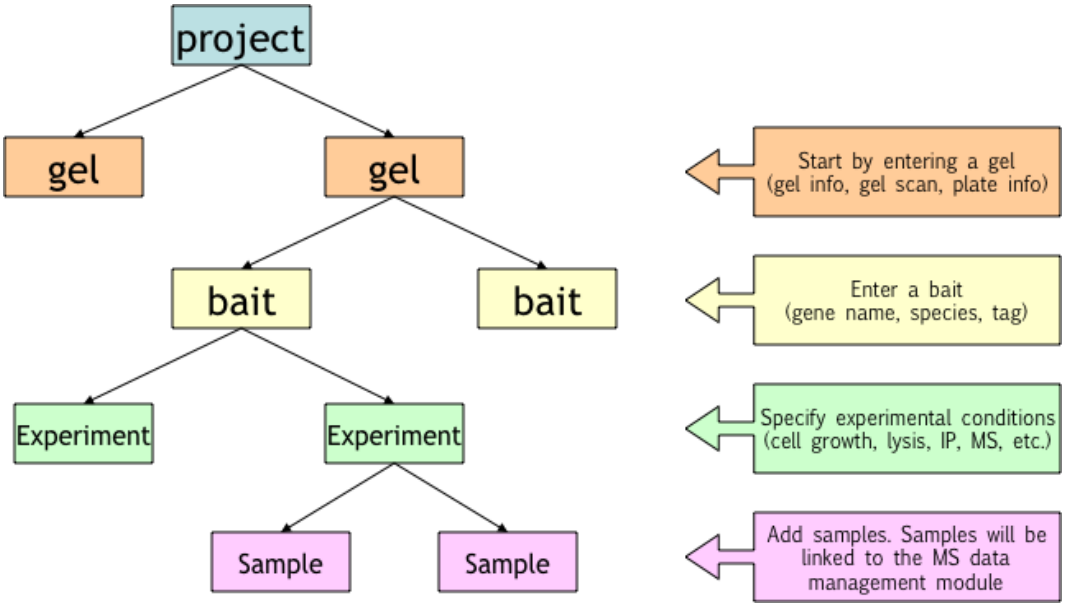
The Analyst module allows you to visualise, analyze, compare, search and export your MS results.



In order to analyze and compare data, each MS file in the MS data management module must be linked to a sample created in the Analyst module. For example, to create a sample for a gel-free experiment, you must first specify a project, create an entry for the protein of interest (bait), and define experimental conditions. Typical gel-free samples are eluates from an affinity purification.



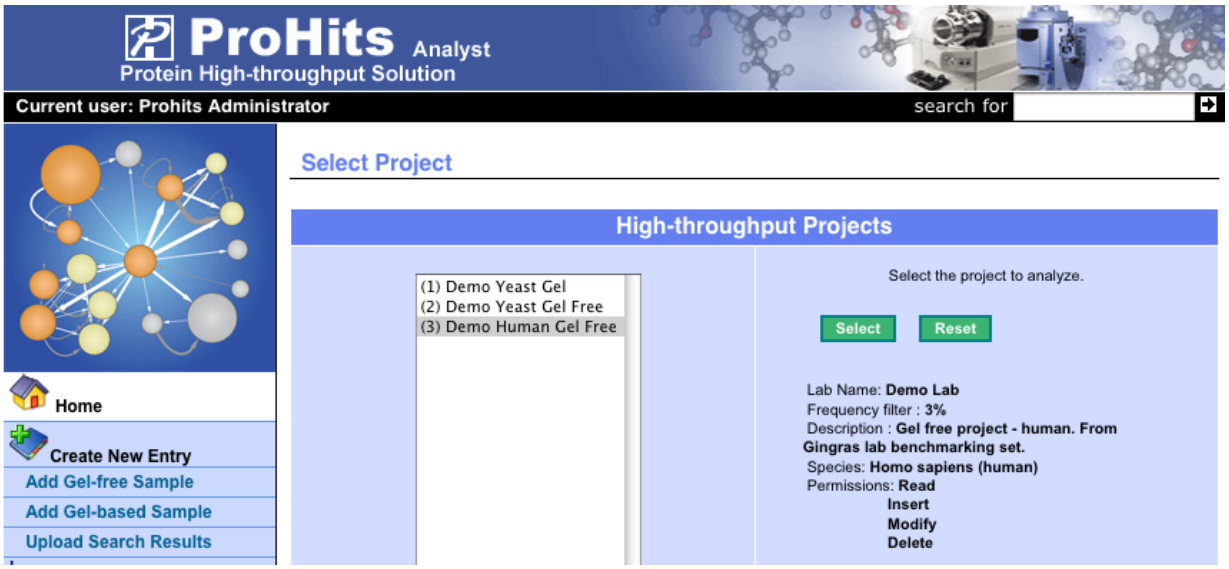
Sample entry for gel-based projects is similar, with the exception that a gel is specified prior to the selection of a bait. Typically, samples are gel bands, and all bands from the same lane are entered under the same "Experiment".



Access to projects
Projects are created by your administrator in the "Admin Office" ProHits module, and access is granted to users. Projects can be specific to a research group or an individual, to a given organism or specific methodology, etc. The creation of a new project is defined in the "Admin Office" manual.

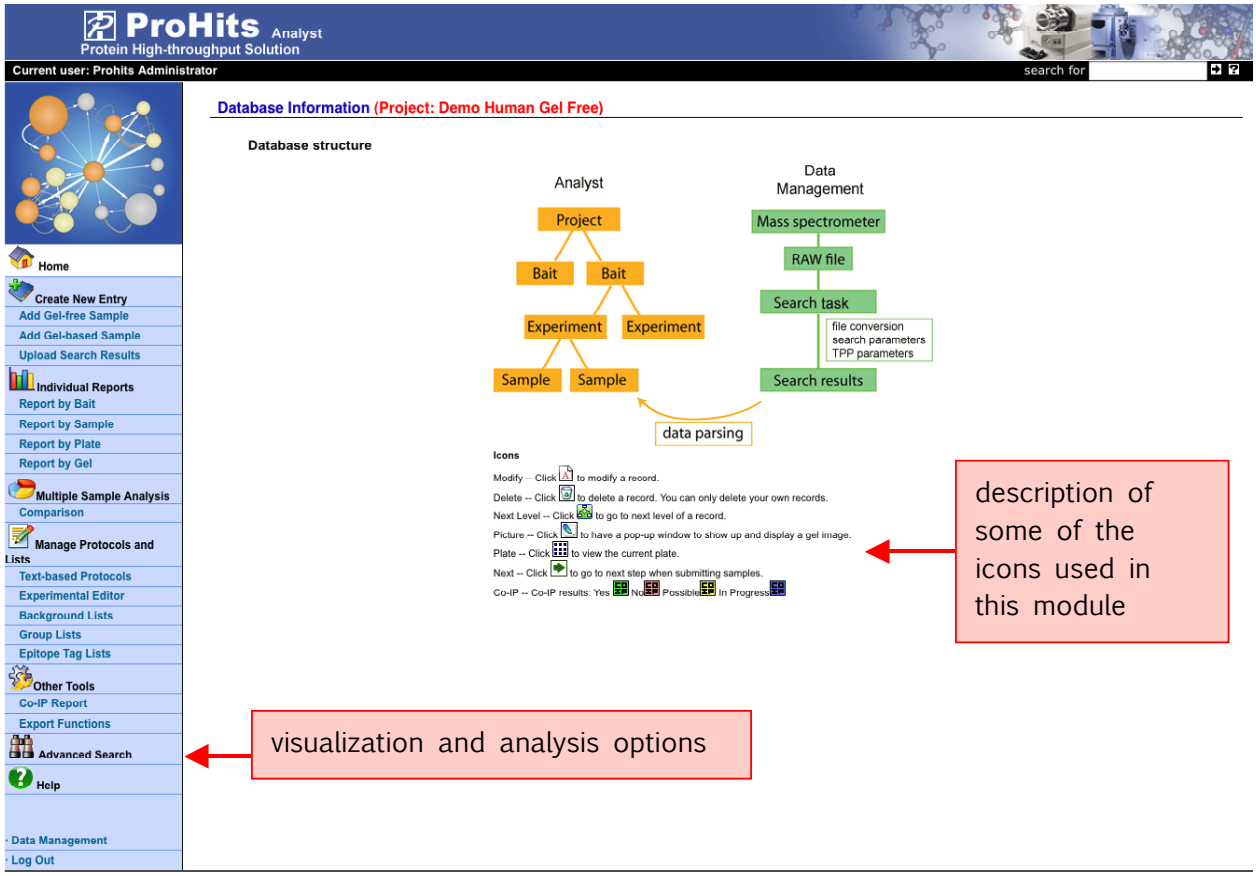
When you log into ProHits with your user name, you can see the list of all of the projects that you have access to. You may have different privileges for each project.

⇒ Highlight the desired project, then hit "Select"



Analyst main page

When you enter a project within the Analyst module, you will see the data workflow and a summary of the icons used in this module. The navigator bar on the left lists various visualization and analysis options.



Description of the navigator bar options:

1- **Create New Entry** allows you to define a bait, experiment, sample, and to link mass spectrometry data to this entry. These entries can then be linked to specific files in the MS Data Management module. Alternatively, you can upload search results created by external software.

2- **Individual Reports** allows you to explore your mass spectrometry results. *Report by Bait*: provides a list of all baits entered in the database for this project. *Report by Samples*: lists all samples entered for this project (a bait may be linked to multiple samples, especially in gel-based projects; we also use this nomenclature for technical replicates). *Report by Plate*: sample tracking for high-throughput projects, typically gel-based. *Report by Gel*: allows you to visualize results for each gel (gel-based projects only).

3- **Multiple Sample Analysis** (Comparison): allows you to simultaneously visualize multiple result pages.

4- **Manage Protocols and Lists** allows you to create and maintain experimental protocols, controlled vocabularies, background lists, group lists and epitope tag lists. Access to these pages is defined by the ProHits Administrator.

5- **Other Tools** provides additional functionality. *Co-IP Report*: allows you to input results from follow-up experiments aimed at confirming interaction pairs by immunoprecipitation/immunoblotting. *Export Functions*: allows you to export filtered or unfiltered lists of mass spectrometry results. Note that export functions are also available within each of the Individual Report or Comparison pages.

6- **Advanced search** allows you to query your project for genes, keywords and/or controlled vocabularies.

Creating samples and viewing individual reports

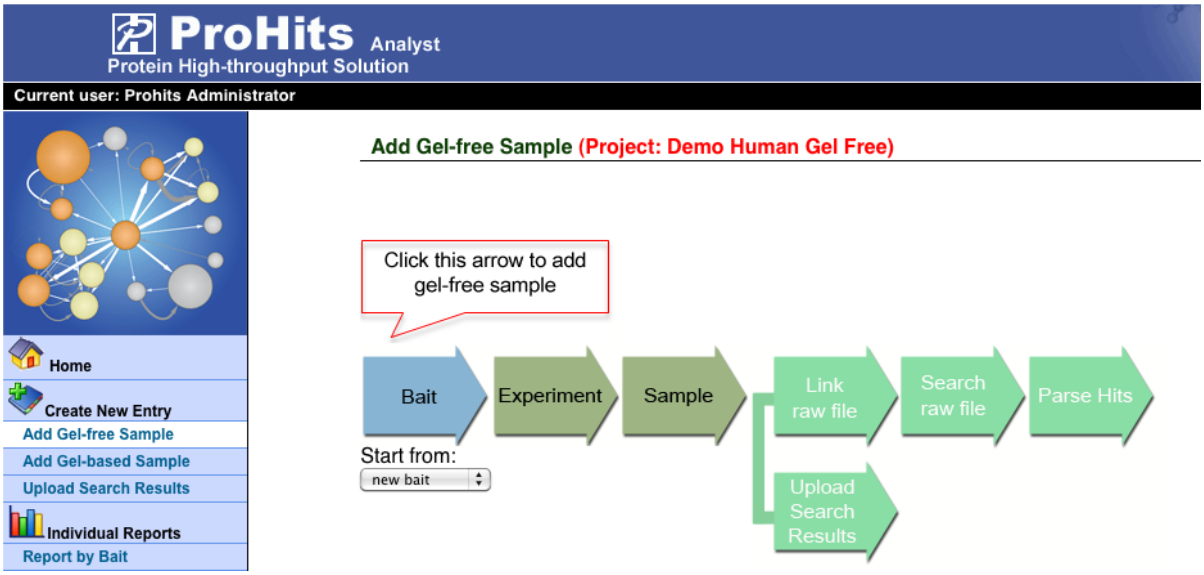
To learn more about the different functions of ProHits Analyst, we will navigate through the Analyst module by creating new baits and linking them to entries from the Data management system. We will then explore the functions available in the Analyst module. We will go through the process of adding a gel-free sample and explore the results for this type of project. We will then briefly review the differences between submitting gel-free and gel-based samples.

Adding a "Gel-Free" sample

To create a new sample to be linked to a search result file, you will first specify a bait, then an experiment, and then a sample. To submit a sample, you have two options: 1) create a new sample from an existing bait; or 2) create a new bait. Here we will start by creating 5 new baits for this project.

Creating a bait

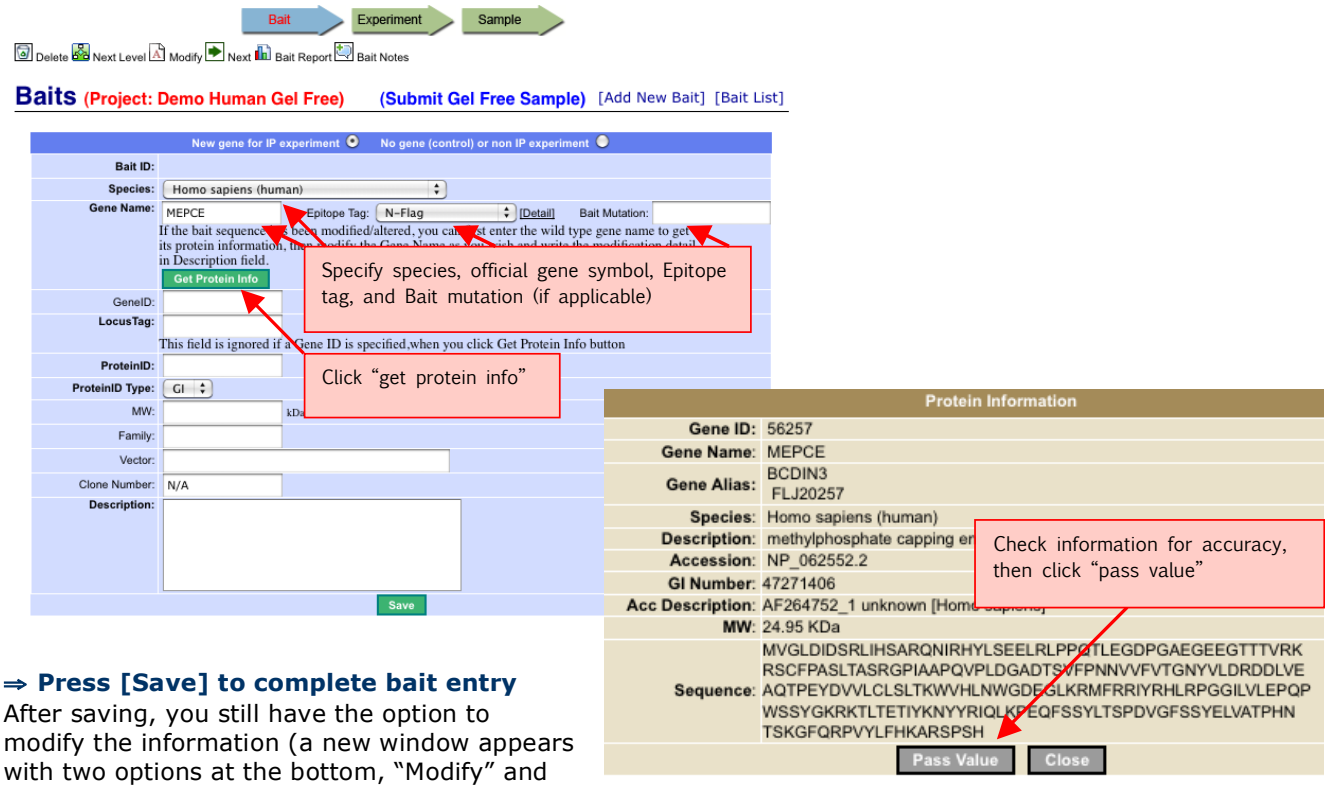
⇒ Select the "Add Gel-free Sample" link under "Create New Entry". Select "new bait" from the dropdown menu, then click on the "Bait" Blue arrow.



This will open a new page. Note at the top of the page the data structure; the Bait is highlighted, indicating that you are adding bait level entries. Note that each of the baits is automatically assigned a unique numeric identifier. The fields highlighted in bold indicate that the information is mandatory, but many of these can be filled automatically.

The easiest way to enter a new bait is to simply 1) select the desired species (here we have selected *Homo sapiens*); 2) enter an official Gene Name (HUGO for human; here we selected MEPCE); 3) click the "Get Protein Info" green button. Clicking "Get Protein Info" automatically retrieves the protein information which is displayed in a new window. Verify this information and hit [Pass Value] if correct – the information will automatically be transferred. Note that if there is more than one entry mapped to a given gene, the user can select which one is to be parsed into ProHits.

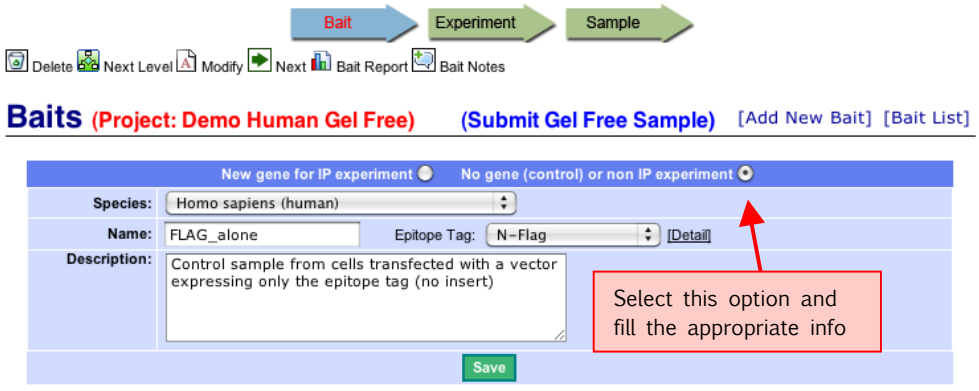
ProHits also allows you to indicate which epitope-tag you are using, by selecting from options in the "Epitope Tag" menu; you can also add new tags using the "Manage Protocols and Lists" option. If the sequence of the bait is mutated relative to the HUGO sequence, you can also enter this in the "Bait mutation" box.



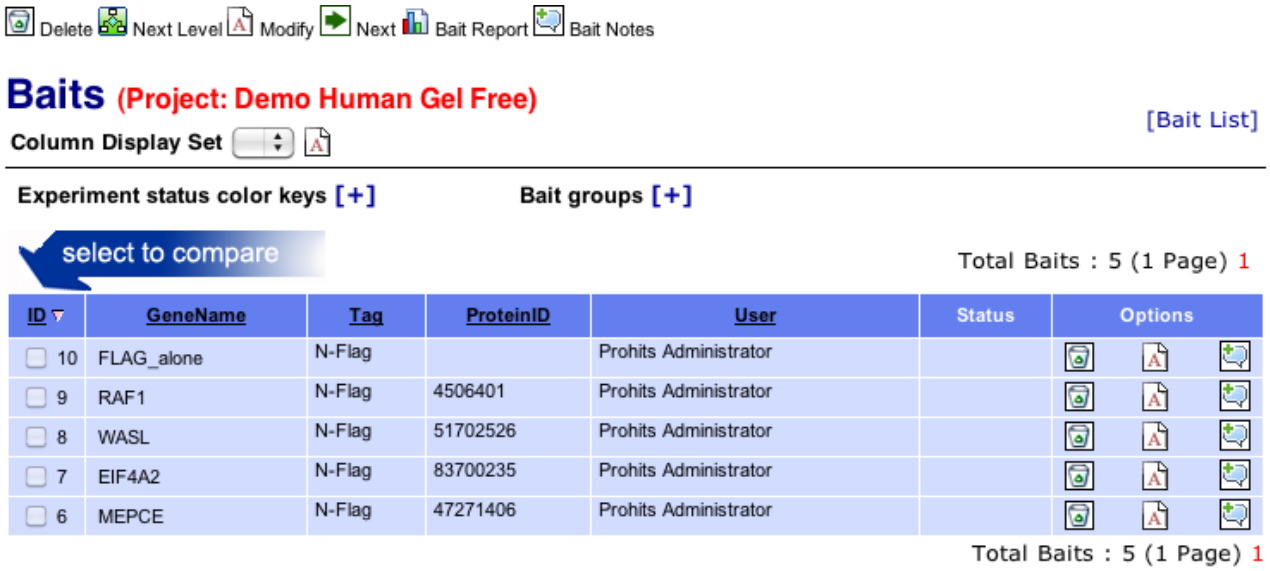
⇒ Press [Save] to complete bait entry
After saving, you still have the option to modify the information (a new window appears with two options at the bottom, "Modify" and "Next"). You can add additional information, e.g. in the "Description" field, or modify existing information. Hitting [Next] would bring up the Experimental detail page (for this demonstration, we will not do this yet).

Also note that you can create baits for sequences that are not in the database by manually filling in all bold fields (species, gene name, locus tag, protein ID, protein ID type). ProHits does not check for accuracy in these entries. You may wish to use this option, for example, for recombinant or chimeric proteins not corresponding to any of the entries in the database.

⇒ Use the [Add New Bait] button at the top of the page, and continue defining baits in the same manner as for MEPCE. Note that in the bait entry page, you can also define an experiment in which no gene/protein was tagged. To do so, simply select the "No gene (control) or non IP experiment" button at the top of the page, and manually enter information. Here we are adding a "FLAG alone" bait.



⇒ To visualize the entry of your new baits in the database, go back to the left bar menu and select [Report by Bait]



The Bait Report now lists the baits we have created (MEPCE, FLAG alone, and 3 additional baits that we will use for the demonstration of the functions of ProHits), along with some relevant information. The "ID" column lists a unique identifier for the bait that is automatically assigned by ProHits. The Gene Name and Tag are indicated, and the Protein ID is the accession number from the selected database (e.g. NCBI-GI). The "User" column is automatically assigned to the user who created the sample (i.e. the person who has signed up in ProHits).

Note that, on many of the ProHits pages, you will find standard icons (as seen at the top of the Bait Report page).

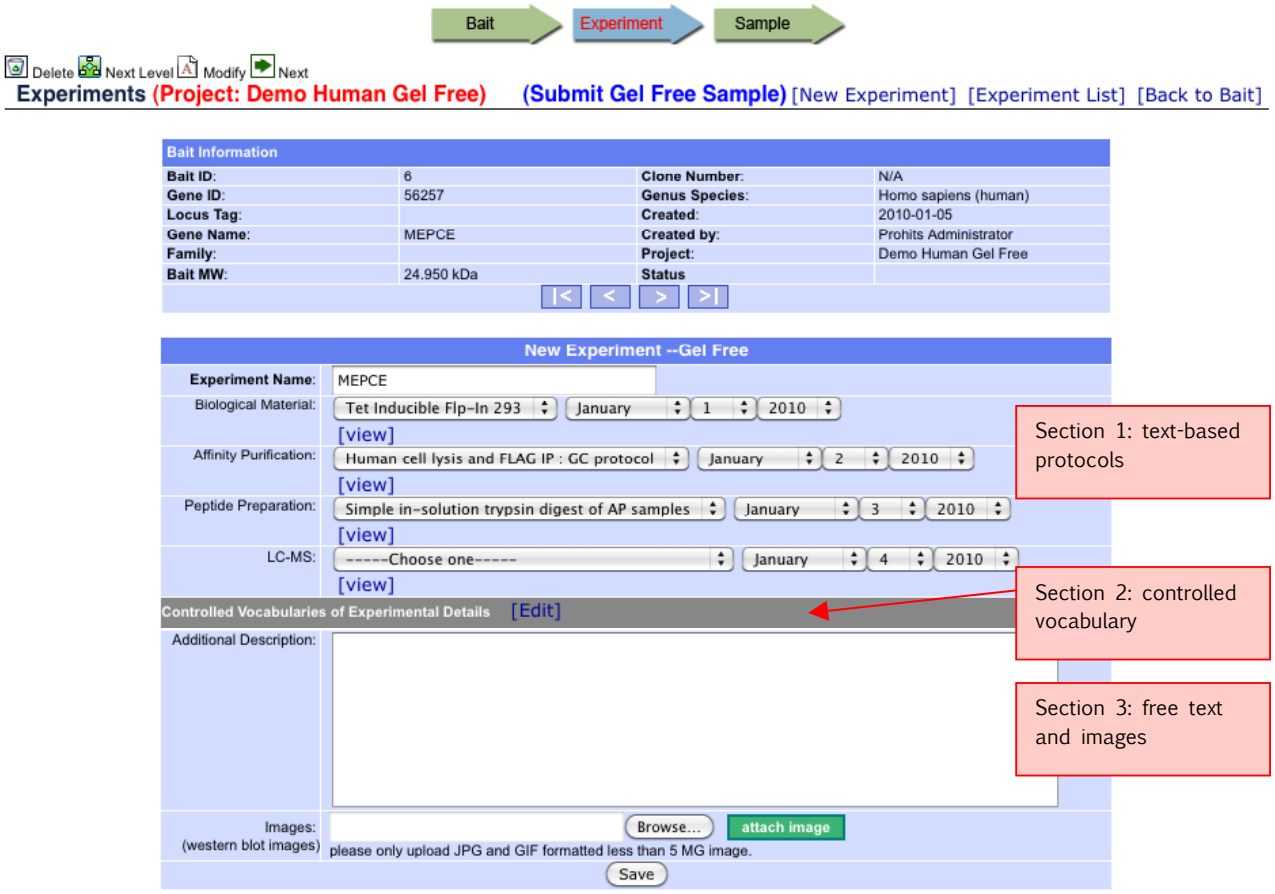
1- To remove unused material, press the "delete" icon. The "delete" function may be used to remove baits, experiments or samples, but *only if no information has been entered*. If you wish to delete a bait, experiment or sample for which information has been entered, start by deleting the information at a lower level, and work your way up. (Note that there is an Admin control for the permissions to insert, modify and delete entries, and you can only delete your own entries).
2- The next level (tree) icon allows you to navigate down in the data structure (i.e. from bait to experiment to sample).
3- The Modify icon allows you to change the information you entered for a bait, experiment or sample.
4- The green arrow (Next) icon allows you to submit information and/or exit a page after data has been entered.
5- The "Bait report" (graph) icon shows you the mass spectrometry results for the selected bait. We will review this in detail later.
6- Finally, the "Bait Notes" (callout) icon allows you to enter specific notes/information for baits or samples. Such notes can be a manually entered discussion point. Other types of notes include assignment of a project to a user-defined "bait group".

Now that you have created baits, you are ready to define your experiments. Note that in many cases, you will be seamlessly going from bait to experiment to sample when entering real samples. Here, we have simply separated these modules for ease of teaching.

Creating an experiment

⇒ Return to [Add Gel-free sample], and select [start from existing bait].
This will bring up essentially the same page as shown above, but with an additional option (green arrow) at the extreme right of each row.

⇒ Select this green arrow to enter the experimental details for a given bait
The experimental detail page allows you to specify experimental conditions and protocols used for the experiment. The top of the page states the bait information: below, the definition of an experiment can be separated into three sections.



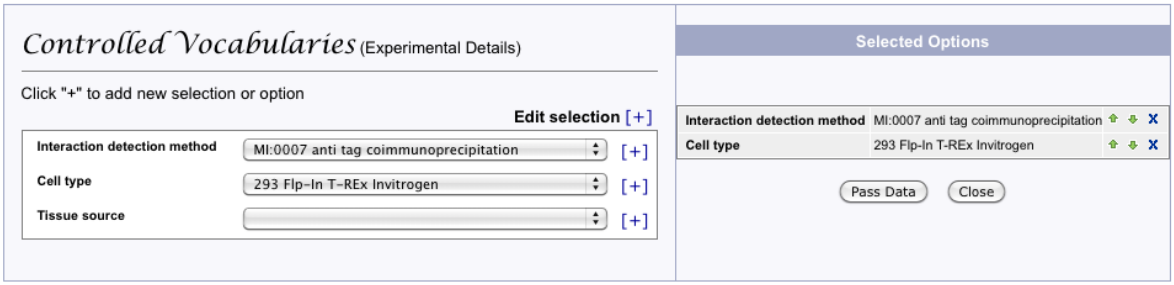
User-defined free-text protocols
In section 1, drop-down menus allow for the selection of user-specified protocols for each experiment. We suggest describing generic protocols in detail (in a manner similar to the Methods section of an article). The protocols can be entered and managed using the "Manage Protocols and Lists" option (more on this later).

Controlled vocabulary
Section 2 offers (via the Experimental Detail Editor) the possibility to specify controlled vocabulary to describe the experiment. The controlled vocabulary is specified for each project by using the "Experimental Editor" option within "Manage Protocols and Lists". Note that this vocabulary can facilitate compliance to community guidelines, such as HUPO Proteomics Standard Initiative (e.g. PSI-MI 2.5). This controlled vocabulary (drop-down keywords) can be used for searching and structuring the data using the "Advanced Search" option.

Additional annotation
Section 3 allows for additional free-text annotation in the form of notes. Here you can cross-reference to notebook page numbers, add specifics of the experiment not captured in sections 1 and 2, or describe any problem or deviation from the reference protocols. It also allows you to link image files (e.g. Western blots or silver stained gels).

⇒ Navigate through the dropdown menus to select appropriate protocols associated with the experiment.

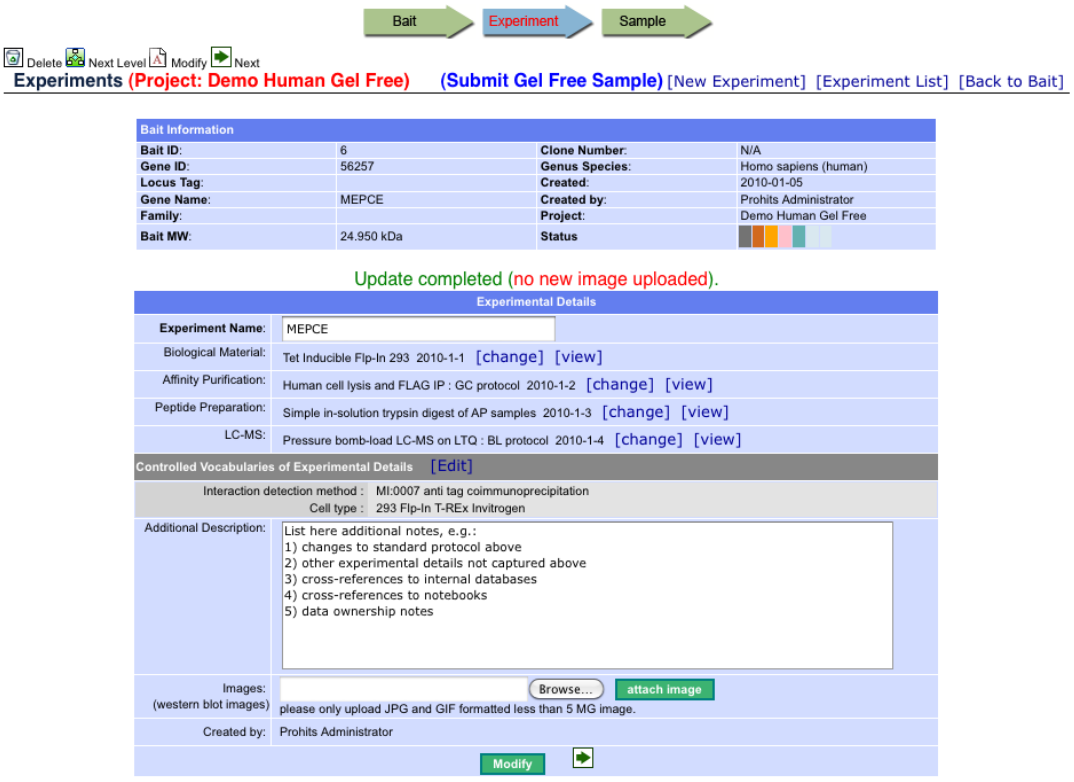
Note that selecting the option "Edit" within Section 2: Controlled Vocabularies of Experimental Details will open up a new window with dropdown menus.



⇒ Select all desired fields to capture using the dropdown menus.
The selected options will be displayed on the right hand side in the order that they were selected. Use the Up/Down green arrows to change the order, or click on the **x** to remove the entry.

⇒ Select [Pass Data] to transfer selection to the Experimental Detail page or [Close] to exit without saving the data.

⇒ Continue filling experimental details, link any desired image, and press [Save].


Bait
Experiment
Sample
Delete Next Level Modify Next
Experiments (Project: Demo Human Gel Free)
(Submit Gel Free Sample) [New Experiment] [Experiment List] [Back to Bait]
Bait Information
Bait ID: 6
Clone Number: N/A
Gene ID: 56257
Genus Species: Homo sapiens (human)
Locus Tag:
Created: 2010-01-05
Gene Name: MEPCE
Created by: Prohits Administrator
Family:
Project: Demo Human Gel Free
Bait MW: 24.950 kDa
Status
Update completed (no new image uploaded).
Experimental Details
Experiment Name: MEPCE
Biological Material: Tet Inducible Flp-In 293 2010-1-1 [change] [view]
Affinity Purification: Human cell lysis and FLAG IP : GC protocol 2010-1-2 [change] [view]
Peptide Preparation: Simple in-solution trypsin digest of AP samples 2010-1-3 [change] [view]
LC-MS: Pressure bomb-load LC-MS on LTQ : BL protocol 2010-1-4 [change] [view]
Controlled Vocabularies of Experimental Details [Edit]
Interaction detection method : MI:0007 anti tag coimmunoprecipitation
Cell type : 293 Flp-In T-REx Invitrogen
Additional Description:
List here additional notes, e.g.:
1) changes to standard protocol above
2) other experimental details not captured above
3) cross-references to internal databases
4) cross-references to notebooks
5) data ownership notes
Images: (western blot images)
Browse...
attach image
please only upload JPG and GIF formatted less than 5 MG image.
Created by: Prohits Administrator
Modify

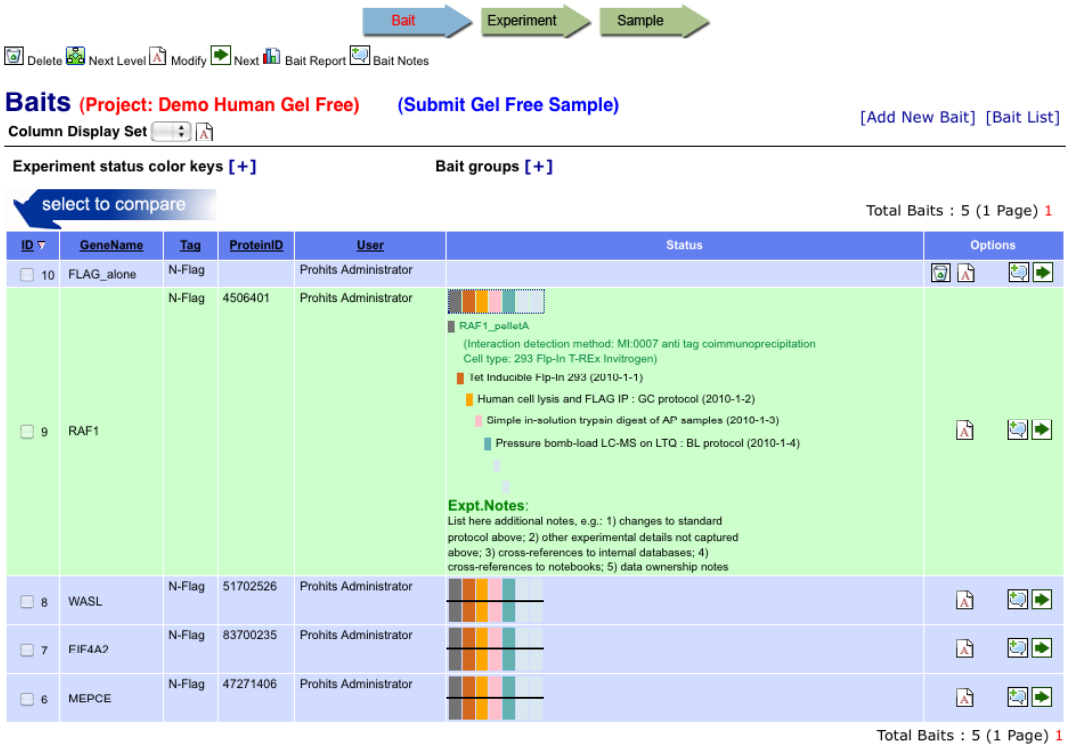
Upon saving, you will be given the option to "Modify" the entry or follow the green arrow to the next page to enter specific samples. Additionally, you can continue creating experiments by toggling between the [New Experiment], [Experiment List] and [Back to Bait] buttons at the top of the page to enter biological replicates for each of the baits.

⇒ Return periodically to the [Back to Bait] list to monitor your progress.

Note the colour-coded experimental status bars in the table. This view shows our five baits, with experiments defined for four of them (MEPCE, EIF4A2, WASL and RAF1). The status column displays experimental details, experimental status and bait groups (see below). The colour-coding in the "Status" column indicates that information has been entered for each of the specified fields.

⇒ Click on the colour-coded status bar to obtain additional experimental details

In the Bait view, experiments (and samples) defined under the same bait will be combined in the same row; multiple experiments will be shown by stacked colour bars. Note that you cannot delete baits for which experiments have been defined (note in the picture below that the FLAG_alone bait can still be deleted, since no experimental details have been entered yet). Start by deleting the Experimental Details, and work your way up as previously described.

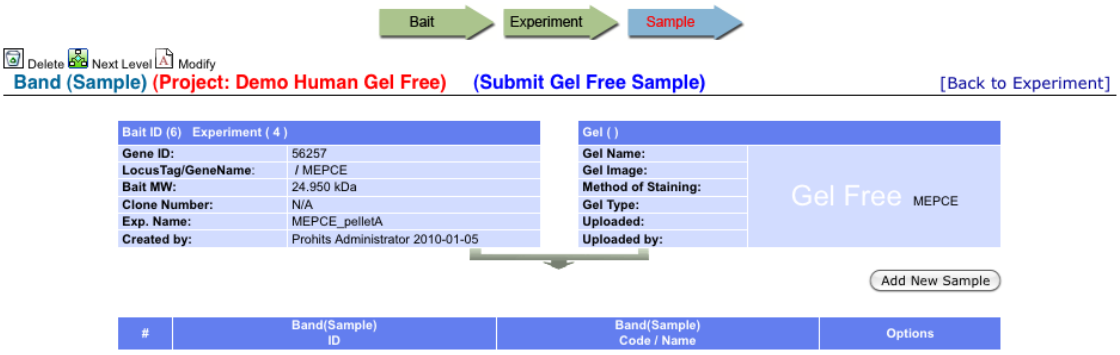

Bait
Experiment
Sample
Delete
Next Level
Modify
Next
Bait Report
Bait Notes
Baits (Project: Demo Human Gel Free)
(Submit Gel Free Sample)
[Add New Bait] [Bait List]
Column Display Set
Experiment status color keys [+]
Bait groups [+]
select to compare
Total Baits : 5 (1 Page) 1
ID
GeneName
Tag
ProteinID
User
Status
Options
10 FLAG_alone N-Flag Prohits Administrator
9 RAF1 N-Flag 4506401 Prohits Administrator
RAF1_pelletA
(Interaction detection method: MI:0007 anti tag coimmunoprecipitation
Cell type: 293 Flp-In T-REx Invitrogen)
Tet Inducible Flp-In 293 (2010-1-1)
Human cell lysis and FLAG IP : GC protocol (2010-1-2)
Simple in-solution trypsin digest of AP samples (2010-1-3)
Pressure bomb-load LC-MS on LTQ : BL protocol (2010-1-4)
Expt.Notes:
List here additional notes, e.g.: 1) changes to standard protocol above; 2) other experimental details not captured above; 3) cross-references to internal databases; 4) cross-references to notebooks; 5) data ownership notes
8 WASL N-Flag 51702526 Prohits Administrator
7 EIF4A2 N-Flag 83700235 Prohits Administrator
6 MEPCE N-Flag 47271406 Prohits Administrator
Total Baits : 5 (1 Page) 1

Once your baits and experiments are entered, you can create one or multiple samples to be linked to the bait and experiment. The number of samples you create for a given experiment depends on your experimental set-up. We tend to use different samples from a single experiment to represent technical replicates (i.e. different MS runs from the same biological sample), where all conditions are the same. Alternatively, multiple samples from one experiment may be created when the sample has been fractionated (e.g. by strong cation exchange) prior to the analysis. Each of the fractions is then assigned a different sample name within the same experiment. The "Notes" sections from the Experimental Details page should explain the sample-naming scheme. Note that we enter biological replicates as different experiments from the same bait.

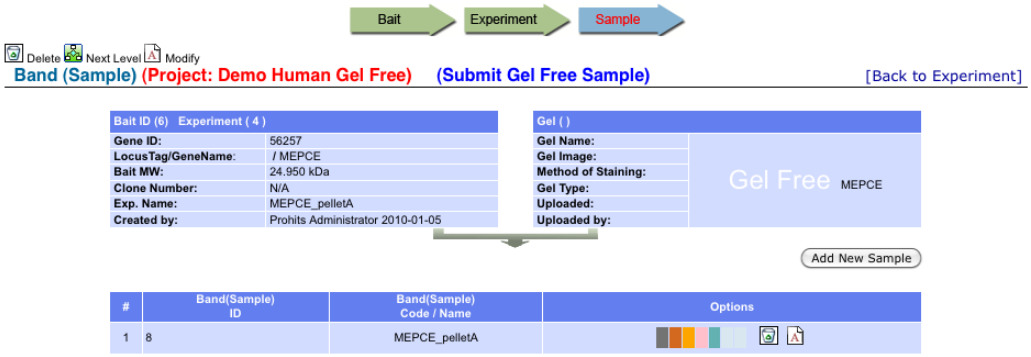
Creating a new sample

Following the green arrow on any of the Experimental Details pages will open a new window, allowing you to create one or many samples for a given experiment.

⇒ In the Sample page, select the [Add New Sample] button to create a sample entry for this bait and set of experimental conditions.



By default, ProHits will use the experiment name to name the first sample created from the relevant experiment. ProHits will also assign a unique Sample ID. The sample name can be modified if necessary (in this case, just type the desired sample name in the text box). In our group, we reserve the creation of duplicate samples from the same bait/experiment for technical replicates (e.g. if we split the final sample in half, and run each half separately). Note that creating multiple samples from a single bait/experiment results in an automatic appending of _A, _B, etc. at the end of the sample name. As long as a sample is not linked to any RAW file, it can be deleted by the owner.



Now that you have created a new sample entry, you are ready to link it to a mass spectrometry raw data file from the Data Management module. Links can be created automatically if the nomenclature indicated in the notice below for file naming is respected, and ProHits Data Management module is connected to the acquisition computers. Alternatively, links can be created manually either from the Data Management or the Analyst modules.

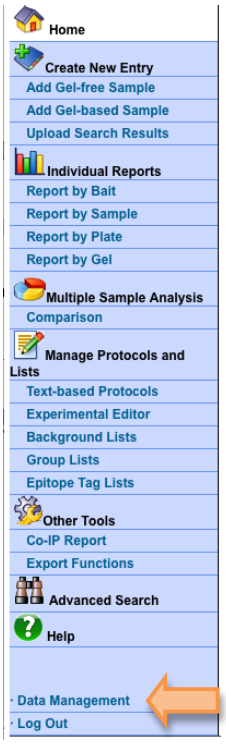
Notice:
In order to link a raw file to a gel free sample automatically, name the folder and raw file as follows.

- raw file(s) is in a folder, the folder name ends with the Project ID.
- start with sample ID and first 4 characters of sample code.

AnyNameNoSpace_P3 (project ID is 3)
8_MEPC_AnyOtherWordNoSpace.RAW

Linking raw files to a created sample

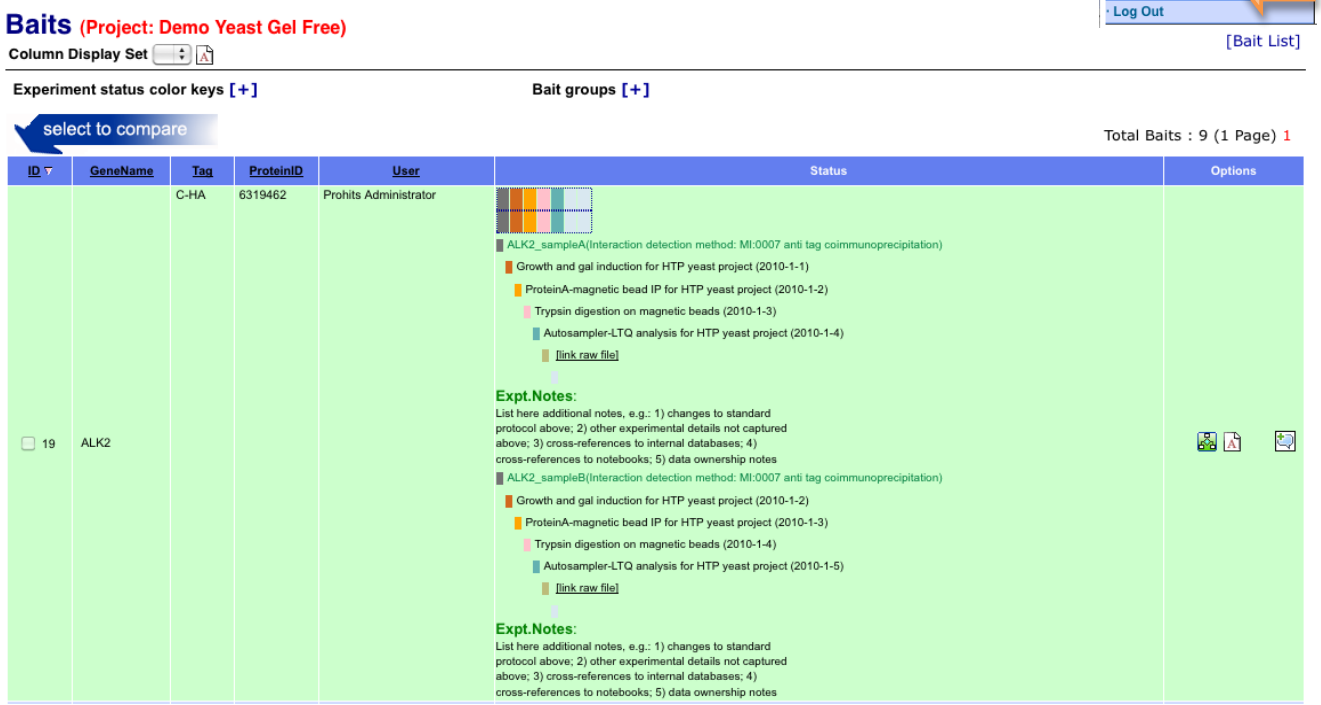
Linking raw files from the Data Management module



⇒ From any page in the Analyst module, click "Data Management" on the left menu bar (shown by orange arrow on the right), link the desired file (as described in the Data Management section), parse the hits and return to the Analyst module.

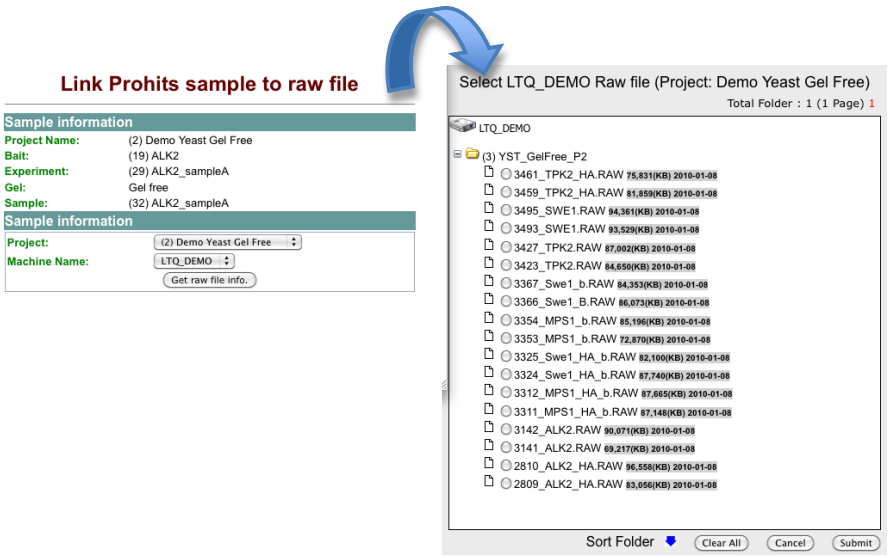
Linking raw files directly from the Analyst module (alternative)
For this alternative example, we are linking files from the Demo Yeast Gel free project, which you can access by going back to the home page of the Analyst module.

⇒ Go to the "Report by Bait" or "Report by Sample" page of the Analyst module and click on the Status column of the desired file to display experimental details. Select [Link raw file].


Baits (Project: Demo Yeast Gel Free)
Column Display Set
[Bait List]
Experiment status color keys [+]
Bait groups [+]
select to compare
Total Baits : 9 (1 Page) 1
ID
GeneName
Tag
ProteinID
User
Status
Options
C-HA
6319462
Prohits Administrator
ALK2_sampleA(Interaction detection method: MI:0007 anti tag coimmunoprecipitation)
Growth and gal induction for HTP yeast project (2010-1-1)
ProteinA-magnetic bead IP for HTP yeast project (2010-1-2)
Trypsin digestion on magnetic beads (2010-1-3)
Autosampler-LTQ analysis for HTP yeast project (2010-1-4)
[link raw file]
Expt.Notes:
List here additional notes, e.g.: 1) changes to standard protocol above; 2) other experimental details not captured above; 3) cross-references to internal databases; 4) cross-references to notebooks; 5) data ownership notes
19
ALK2
ALK2_sampleB(Interaction detection method: MI:0007 anti tag coimmunoprecipitation)
Growth and gal induction for HTP yeast project (2010-1-2)
ProteinA-magnetic bead IP for HTP yeast project (2010-1-3)
Trypsin digestion on magnetic beads (2010-1-4)
Autosampler-LTQ analysis for HTP yeast project (2010-1-5)
[link raw file]
Expt.Notes:
List here additional notes, e.g.: 1) changes to standard protocol above; 2) other experimental details not captured above; 3) cross-references to internal databases; 4) cross-references to notebooks; 5) data ownership notes

This brings up a new page that allows you to select the file to be linked to the given entry.

Note that when you link files from the Analyst module, only those files not previously linked to another entry will be displayed. To modify an existing link, you need to go back to the Data Management module, remove the link to the initial file, so that it can be made available to link to an entry either through the Analyst or Data Management modules.

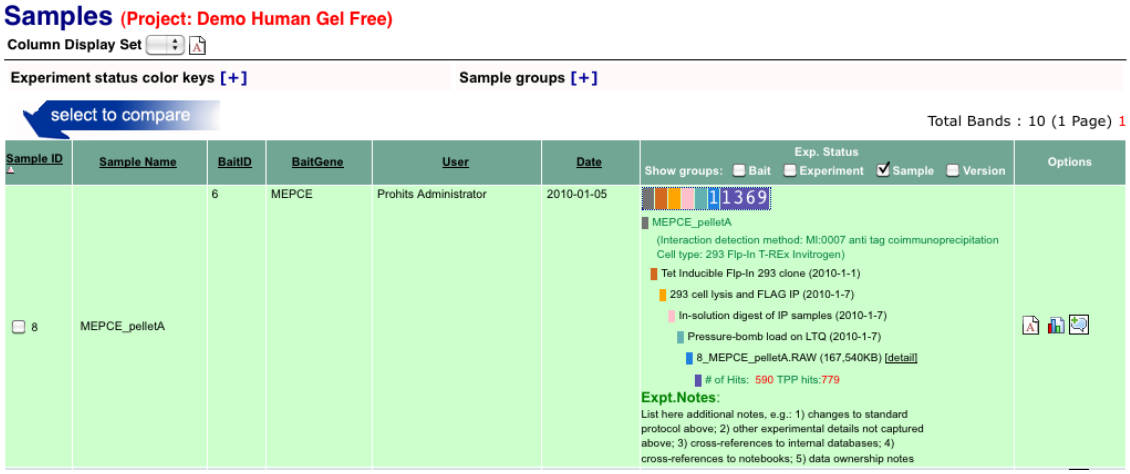


Once a raw file has been linked, the status bar will display an additional blue icon; the number indicates the number of files linked to that entry.

Baits (Project: Demo Yeast Gel Free)
Column Display Set
[Bait List]
Experiment status color keys [+]
Bait groups [+]
select to compare
Total Baits : 9 (1 Page) 1
ID
GeneName
Tag
ProteinID
User
Status
Options
C-HA
6319462
Prohits Administrator
ALK2_sampleA(Interaction detection method: MI:0007 anti tag coimmunoprecipitation)
Growth and gal induction for HTP yeast project (2010-1-1)
ProteinA-magnetic bead IP for HTP yeast project (2010-1-2)
Trypsin digestion on magnetic beads (2010-1-3)
Autosampler-LTQ analysis for HTP yeast project (2010-1-4)
2809_ALK2_HA.RAW (83,056KB) [detail]
Expt.Notes:
List here additional notes, e.g.: 1) changes to standard protocol above; 2) other experimental details not captured above; 3) cross-references to internal databases; 4) cross-references to notebooks; 5) data ownership notes
19
ALK2
ALK2_sampleB(Interaction detection method: MI:0007 anti tag coimmunoprecipitation)
Growth and gal induction for HTP yeast project (2010-1-2)
ProteinA-magnetic bead IP for HTP yeast project (2010-1-3)
Trypsin digestion on magnetic beads (2010-1-4)
Autosampler-LTQ analysis for HTP yeast project (2010-1-5)
[link raw file]
Expt.Notes:
List here additional notes, e.g.: 1) changes to standard protocol above; 2) other experimental details not captured above; 3) cross-references to internal databases; 4) cross-references to notebooks; 5) data ownership notes

⇒ Click the "Data Management" tab from any page of the Analyst module, parse the hits (as described in the Data Management section) and return to the Analyst module.

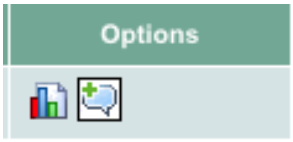
Once hits are parsed (either from the Data Management or the Analyst module), a new purple coloured tab will appear in the status bar (in either Bait Report or Sample Report pages), indicating the total number of hits identified (sum of hits if more than one search engine was used). In the "Options" column, a new graph icon appears; clicking this link brings up the search results for each sample. Here we are showing MEPCE_pellet A in the sample report view.



You are now ready to explore your results. Use the left-hand side of the ProHits Analyst main page to view "Report by Bait" and "Report by Sample". The interface for the Bait and Sample reports is very similar. Here we provide an example for the Sample Report. *Bait versus Sample view*: For some projects you may have a one-to-one correspondence between bait and sample. For other projects, you will have multiple samples linked to the same bait. Opening the Bait Report when two or more samples are linked to the bait will generate sequential protein hit lists for each of the samples linked to the bait. ProHits does not recalculate scores or peptide numbers, but indicates (in bold) proteins detected in more than one sample (mousing over bolded names activates a pop-up window that provides details about the samples and hit scores). If you wish to explore each sample individually, use the "Report by Sample" link instead.

Navigating through the results

Now that we have entered baits, linked and parsed search results, it is time to look at search results. In this example, we will start from the "Report by Sample" page for MEPCE_pelletA.



⇒ From the sample list page, under "Options", select the graph icon from one of the samples to see the results.

The following page appears, displaying the results from your search engine (Mascot in this example), alongside links to initial search results and biological databases. Additional export and viewing functions, as well as options to filter the hits are also available from this page. Over the next several pages, we will explore the Results page.

Exclusion Color
Gel image
Mascot
GPM
Hit Notes
Yes
No
Possible
In Progress
Sample Report Hits
(Project: Demo Human Gel Free)
[Cytoscape] [Export Sample Report] [Back to Sample List]
Bait Information (6) [GI][Gene][BioGrid]
Bait Gene ID 56257
Bait Gene Name MEPCE (N-Flag)
Bait Locus Tag
Bait MW (kDa) 24.950
Bait Clone N/A
Bait Description AF264752_1 unknown [Homo sapiens]
Experiment Information
Experiment ID Name/Batch Name Exp. Detail Exp. Status: Inputed by Dat
4 MEPCE_pelletA Interaction detection method : MI:0007 anti tag coimmunoprecipitation; Cell type : 293 Flp-In T-REx Invitrogen 11369 Prohits Administrator 2010-01-05
Sample Information
Sample ID Sample Code Submitted by
8 MEPCE_pelletA Prohits Administrator
Bait description and Experimental details
Export and visualization
Total Hits: 343
335
64
7
1
1
2
NS CO SO AW FQ RP CP BT KT AT TE DB AL
[Show Filters]
Click to enable filtering options
Use tabs to navigate through results pages
Update Frequency
Search results
Mascot Hits
GPM Hits
Mascot TPP Hits
Mascot TPP Peptides
Other TPP Hits
Other TPP Peptides
ID Protein Gene Score Expect Frequency Redundant MW kDa Description # Peptide # Unique Peptide Coverage Links Filter Option
47271406 56257 / MEPCE 1106 20% 74.310 bin3; bicoid-interacting 3 [Homo sapiens] 152 19 43.80% [GI][Gene][BioGrid]
425 4502491 708 / C1QBP 686 100% 31.340 complement component 1; q subcomponent binding protein precursor [Homo sapiens] 138 9 60.60% [GI][Gene][BioGrid]
413 7661952 9733 / SART3 1463 20% 109.860 squamous cell carcinoma antigen recognized by T cells 3 [Homo sapiens] 105 20 26.90% [GI][Gene][BioGrid]
416 55956788 4691 / NCL 1011 100% 76.570 nucleolin [Homo sapiens] 84 16 25.80% [GI][Gene][BioGrid]
430 10835063 4869 / NPM1 657 100% 40353734; 32.550 nucleophosmin 1 isoform 1 [Homo sapiens] 65 7 36.70% [GI][Gene][BioGrid]

Search results

Towards the bottom of the page are the search results – by default, these are not filtered. The red colour in the ID field indicates the bait (as defined by the user when entering the experimental description). There are several tabs at the top of the search results table available for navigation. The exact tabs displayed depend on the search engines used. For the demonstration project, we have used the search engines Mascot and X!Tandem (GPM), and have analysed the results using the TransProteomics Pipeline. We will first explore the "Mascot Hits" tab.

The columns list the following parameters:

A) ID: Unique identifier assigned by ProHits (for database purposes)
B) Protein: Protein accession number from original database used by the search engine
C) Gene: NCBI Gene ID/ Gene Symbol, mapped by ProHits from Protein accession
D) Score: Mascot score (if applicable)
E) Expect value: GPM / X!Tandem Expect value (if applicable)
F) Frequency: The frequency that this protein hit is detected across all samples analyzed for this project
G) Redundant: Other protein accession numbers matching the same set of peptides
H) MW kDa: Calculated MW for the protein
I) Description: Definition field from the NCBI protein entry
J) # Peptide: Spectral counts (or total peptides), as calculated by the search engine
K) # Unique Peptide: Number of unique peptides, as calculated by the search engine

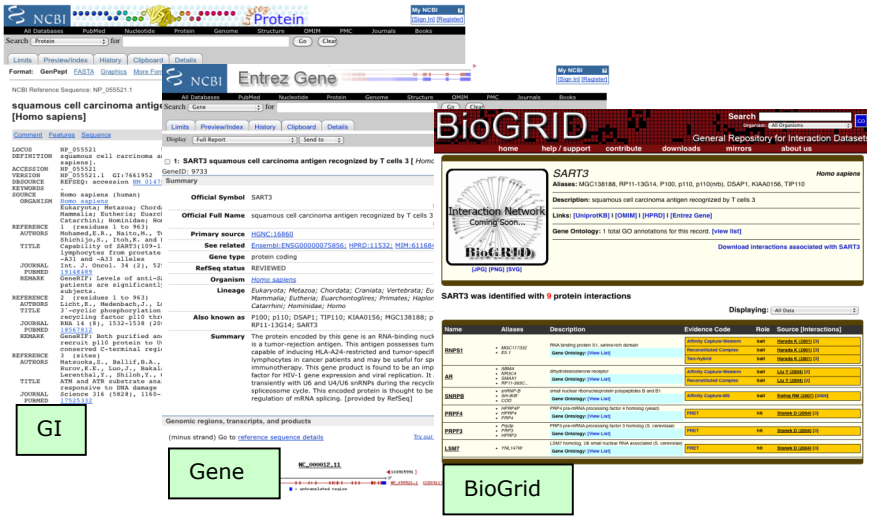
L) Coverage: Percentage of the indicated amino acid sequence identified by your search engine
M) Links: External links to the NCBI Entrez Protein page [GI], the NCBI Gene Page [Gene] and the BioGrid [BioGrid].
N) Filter: provides a colour-coded view of the Experimental Filters or Bio Filters that could be applied to remove each hit
O) Option: Provides the list of peptides belonging to this hit (green M icon), opens up the original search engine search results (here Matrix Science icon for Mascot search results), and allows for the addition of Notes (call-out icons; includes manual exclusion)

Sorting options

You can sort the results from any of the black underlined columns (Score, #Peptide, #UniquePeptide and Coverage); sorting can be in ascending or descending value.

Links details

The following pages can be obtained from each of the items in the "Links" column.



Option details

Pressing the following icons in the Option column will retrieve the peptide list (from the search engine) for each hit, or the entire search results file.

Hit Information (7661952[GI])

Instrument	ESI-TRAP	Score/Expect	1463/
Redundant		Results File	Click to view Mascot search results
Search Database	Human_RefseqV33 Human_RefseqV33_20090226.fasta-All entries	Search Date	2010-01-09 00:33:49

ID	Score	Expect	Charge	Mass (kDa)	Location	Sequence
1914	35	2	1	1.619750	310–322	LAEYQAYIDFEMK
1904	35	2	1	1.620740	310–322	LAEYQAYIDFEMK + 1 Deamidated (NQ),Q5
1971	35	3	2	2.964540	919–950	ALQRPSAAAPQAENGPAAAPAVAAPAATEAPK
1907	35	2	1	1.619750	310–322	LAEYQAYIDFEMK
1903	36	2	1	1.595790	392–405	HGVDHQVISVTFEK + 1 Deamidated (NQ),Q6
1876	36	2	1	1.077560	224–232	GLALWEAYR
1915	37	2	1	1.619750	310–322	LAEYQAYIDFEMK
1928	37	2	1	1.619750	310–322	LAEYQAYIDFEMK
1916	37	2	1	1.619750	310–322	LAEYQAYIDFEMK
1873	38	2	0	0.917530	329–335	IQLIFER
1933	38	2	1	1.619750	310–322	LAEYQAYIDFEMK
1895	38	3	1	1.594810	392–405	HGVDHQVISVTFEK
1963	38	3	2	2.271060	702–721	DSITVFVSNLPYSMQEPDTK + 1 Deamidated (NQ),Q15
1930	38	2	1	1.619750	310–322	LAEYQAYIDFEMK
1911	38	2	1	1.619750	310–322	LAEYQAYIDFEMK
1875	38	2	1	1.077560	224–232	GLALWEAYR
1901	39	3	1	1.594810	392–405	HGVDHQVISVTFEK
1898	39	3	1	1.594810	392–405	HGVDHQVISVTFEK
1946	39	3	1	1.776900	295–309	YKPYEEALLQAEAPR
1934	39	2	1	1.620740	310–322	LAEYQAYIDFEMK + 1 Deamidated (NQ),Q5
1966	39	3	2	2.364200	646–669	RVENSIPAAGETQNVEVAAGPAGK
1921	39	2	1	1.619750	310–322	LAEYQAYIDFEMK

[Download DAT]

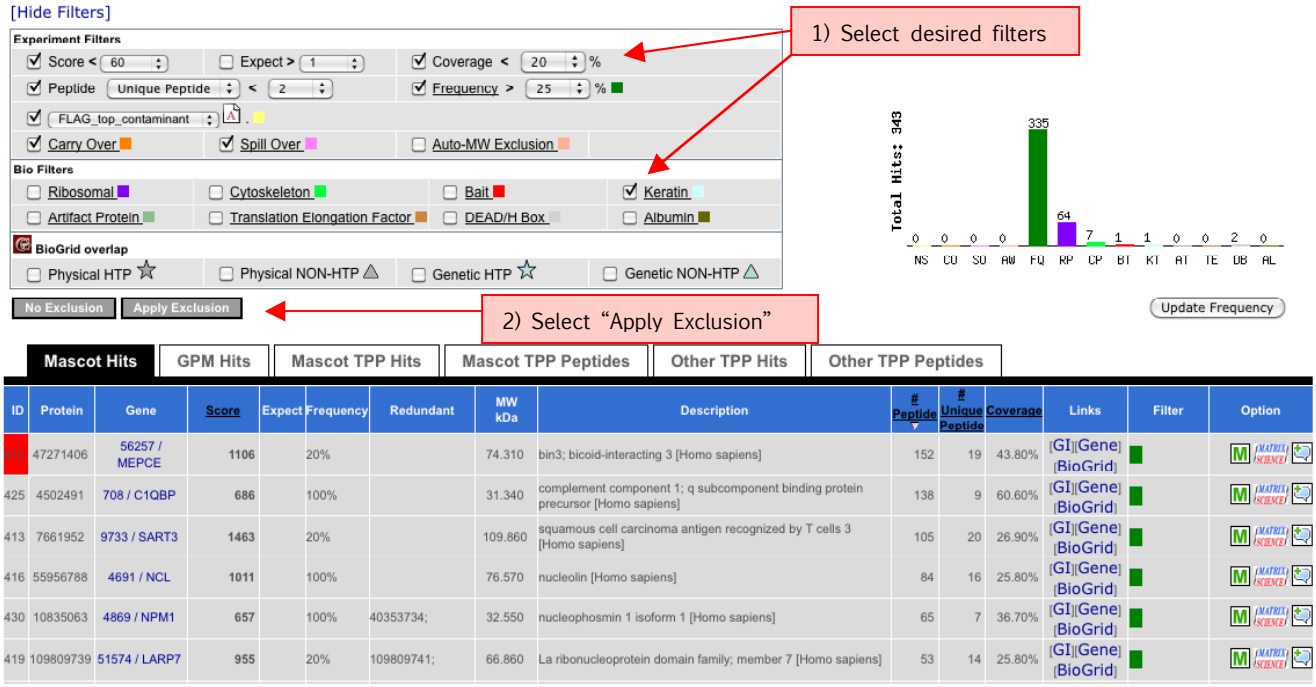
Mascot Search Results

User : frank liu
Email : gliu@mshri.on.ca
Search title : demo human
MS data file : 8_MEPCE_pelletA.mgf
Database : Human_RefseqV33 RefseqV33_20090226 (38930 sequences; 18115101 residues)
Timestamp : 7 Jan 2010 at 18:45:50 GMT
Protein hits : gi|40217867 activating signal cointegrator 1 complex subunit 3-like 1 [Homo sapiens]
gi|91208426 U5 snRNP-specific protein [Homo sapiens]
gi|7661952 squamous cell carcinoma antigen recognized by T cells 3 [Homo sapiens]
gi|13654237 protein kinase, DNA-activated, catalytic polypeptide isoform 1 [Homo sapiens]
gi|47271406 bin3, bicoid-interacting 3 [Homo sapiens]
gi|55956788 nucleolin [Homo sapiens]
gi|217272892 elongation factor Tu GTP binding domain containing 2 isoform a [Homo sapiens]
gi|109809739 La ribonucleoprotein domain family, member 7 [Homo sapiens]
gi|13699824 kinesin family member 11 [Homo sapiens]
gi|100913206 DEAH (Asp-Glu-Ala-His) box polypeptide 9 [Homo sapiens]
gi|19923142 karyopherin beta 1 [Homo sapiens]
gi|10863945 ATP-dependent DNA helicase II [Homo sapiens]
gi|5729877 heat shock 70kDa protein 8 isoform 1 [Homo sapiens]
gi|50659095 DEAD (Asp-Glu-Ala-Asp) box polypeptide 21 [Homo sapiens]
gi|24431950 PRP4 pre-mRNA processing factor 4 homolog [Homo sapiens]
gi|4504447 heterogeneous nuclear ribonucleoprotein A2/B1 isoform A2 [Homo sapiens]
gi|20070220 protein arginine methyltransferase 5 isoform a [Homo sapiens]
gi|4758556 PRP3 pre-mRNA processing factor 3 homolog [Homo sapiens]
gi|10835063 nucleophosmin 1 isoform 1 [Homo sapiens]
gi|4503841 ATP-dependent DNA helicase II, 70 kDa subunit [Homo sapiens]
gi|4502491 complement component 1, q subcomponent binding protein precursor [Homo sapiens]
gi|10863889 squamous cell carcinoma antigen recognized by T cells 1 [Homo sapiens]
gi|221136939 pre-mRNA processing factor 31 homolog [Homo sapiens]
gi|4504897 karyopherin alpha 2 [Homo sapiens]
gi|21397427 synaptotagmin binding, cytoplasmic RNA interacting protein [Homo sapiens]
gi|156523968 poly (ADP-ribose) polymerase family, member 1 [Homo sapiens]
gi|29788785 tubulin, beta [Homo sapiens]
gi|10835067 autoantigen La [Homo sapiens]
gi|14141152 heterogeneous nuclear ribonucleoprotein M isoform a [Homo sapiens]
gi|4506671 ribosomal protein P2 [Homo sapiens]
gi|15718687 ribosomal protein S3 [Homo sapiens]
gi|167466173 heat shock 70kDa protein 1b [Homo sapiens]
gi|5174735 tubulin, beta, 2 [Homo sapiens]
gi|17978466 cyclin T1 [Homo sapiens]
gi|4504445 heterogeneous nuclear ribonucleoprotein A1 isoform a [Homo sapiens]
gi|4502891 chloride channel, nucleotide-sensitive, 1A [Homo sapiens]
gi|157694492 MYB binding protein 1a isoform 2 [Homo sapiens]

We have now navigated through the table listing the search results. However, the initial list is not filtered; that is, all hits, including likely contaminants, are listed. ProHits has a built-in filter set that can be applied to the data to help identify *bona fide* interactors.

Using filters

Click on the [Show Filters] button within the results page to display the administrator-defined Bio and Experimental filters (see admin office for details of the filtering options) and background lists (see Manage Protocols and Lists) that can be applied to the data in this project. On the left is the filter list and the graph on the right indicates the number of proteins that would be removed by activating each of the filters. Filters are activated or de-activated by clicking their associated checkbox. Once the desired filters are selected, press "Apply exclusion" to remove associated proteins from the search results list. Note that the default frequency filter is set in the admin office module when creating the project, and that this value is listed when you select the project from the home page (see page 3). In the case of the "Demo Human Gel Free" project shown here, the frequency filter was set at 3%, meaning that a protein detected in >3% of samples within the project is flagged (as shown by the dark green icon in the results table). You do not need to use the default filter, and can modify this frequency cut-off as needed. Also note that the frequency is not automatically recalculated every time you add a search result to ProHits: to recalculate the frequency, use the "Update Frequency" button on any "Report" page.

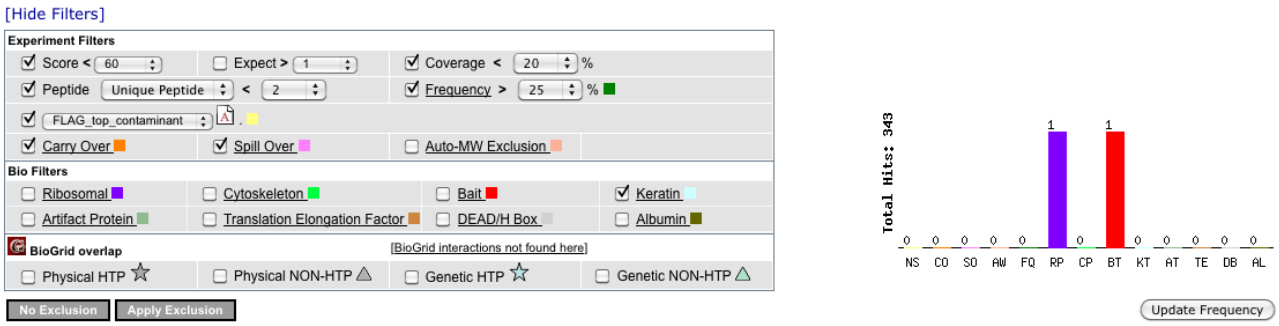


ID	Protein	Gene	Score	Expect	Frequency	Redundant	MW kDa	Description	# Peptide	# Unique Peptide	Coverage	Links	Filter	Option
[redacted]	47271406	56257 / MEPCE	1106		20%		74.310	bin3; bicoid-interacting 3 [Homo sapiens]	152	19	43.80%	[GI][Gene] [BioGrid]		
425	4502491	708 / C1QBP	686		100%		31.340	complement component 1; q subcomponent binding protein precursor [Homo sapiens]	138	9	60.60%	[GI][Gene] [BioGrid]		
413	7661952	9733 / SART3	1463		20%		109.860	squamous cell carcinoma antigen recognized by T cells 3 [Homo sapiens]	105	20	26.90%	[GI][Gene] [BioGrid]		
416	55956788	4691 / NCL	1011		100%		76.570	nucleolin [Homo sapiens]	84	16	25.80%	[GI][Gene] [BioGrid]		
430	10835063	4869 / NPM1	657		100%	40353734;	32.550	nucleophosmin 1 isoform 1 [Homo sapiens]	65	7	36.70%	[GI][Gene] [BioGrid]		
419	109809739	51574 / LARP7	955		20%	109809741;	66.860	La ribonucleoprotein domain family; member 7 [Homo sapiens]	53	14	25.80%	[GI][Gene] [BioGrid]		

In this example, we will filter the data shown above by applying the following filters:

1) click the "background" button, and select the "FLAG_top_contaminants" list from the dropdown menu. The background lists are user-defined, and controlled via the "Manage Lists and Protocols" option.
2) proteins detected with a Mascot score <60 will be removed
3) proteins with <20% sequence coverage will be removed
4) proteins detected with a single unique peptide will be removed
5) proteins detected in >25% of the samples in this project will be removed

After applying filters, the list of hits is reduced (see the disappearance of C1QBP, NCL and NPM1 – which are common contaminants - while SART3, LARP7 and LSM8 remain). The filters can be modified and sorting repeated: ProHits does not remove any data from the dataset, but only displays filtered lists.



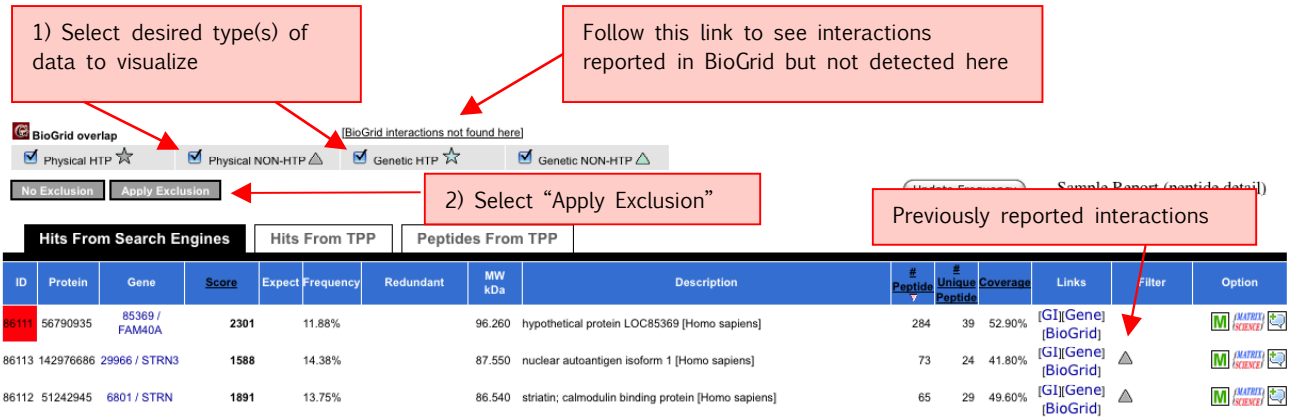
Mascot Hits | GPM Hits | Mascot TPP Hits | Mascot TPP Peptides | Other TPP Hits | Other TPP Peptides

ID	Protein	Gene	Score	Expect	Frequency	Redundant	MW kDa	Description	# Peptide	# Unique Peptide	Coverage	Links	Filter	Option
415	47271406	56257 / MEPCE	1106		20%		74.310	bin3; bicoid-interacting 3 [Homo sapiens]	152	19	43.80%	[GI][Gene] [BioGrid]		
413	7661952	9733 / SART3	1463		20%		109.860	squamous cell carcinoma antigen recognized by T cells 3 [Homo sapiens]	105	20	26.90%	[GI][Gene] [BioGrid]		
419	109809739	51574 / LARP7	955		20%	109809741;	66.860	La ribonucleoprotein domain family; member 7 [Homo sapiens]	53	14	25.80%	[GI][Gene] [BioGrid]		
468	7706425	51691 / LSM8	305		20%		10.400	U6 snRNA-associated Sm-like protein LSm8 [Homo sapiens]	49	3	61.50%	[GI][Gene] [BioGrid]		

Note that the graph on the right indicates the number of hits that have not been filtered out, but belong to the different categories that could be filtered out. In this example, after filtering, only 1 RP (Ribosomal Protein) remains, as compared to 64 in the unfiltered example.

Comparing your data to literature interactions

Prohits allows you to automatically query the BioGrid interaction database for previously-reported interactions specific to your bait. To do so, select the type of interactions desired (physical interactions from high-throughput (HTP) studies, physical interactions not from HTP studies (non-HTP), genetic interactions of both types), and press "Apply exclusion". The interactions that overlap with the literature will be highlighted in the "filter" column. (the next few figures will be replaced by MEPCE as soon as the new version of BioGrid comes online). Note that the definition of HTP and non-HTP is from BioGrid: high-throughput papers are identified as such by BioGrid curators; as a default, publications reporting >100 interactions are also identified as HTP.


1) Select desired type(s) of data to visualize
Follow this link to see interactions reported in BioGrid but not detected here
BioGrid overlap
[BioGrid interactions not found here]
Physical HTP
Physical NON-HTP
Genetic HTP
Genetic NON-HTP
No Exclusion
Apply Exclusion
2) Select "Apply Exclusion"
Previously reported interactions
Hits From Search Engines
Hits From TPP
Peptides From TPP
ID Protein Gene Score Expect Frequency Redundant MW kDa Description # Peptide # Unique Peptide Coverage Links Filter Option
86111 56790935 85369 / FAM40A 2301 11.88% 96.260 hypothetical protein LOC85369 [Homo sapiens] 284 39 52.90% [GI][Gene][BioGrid]
86113 142976686 29966 / STRN3 1588 14.38% 87.550 nuclear autoantigen isoform 1 [Homo sapiens] 73 24 41.80% [GI][Gene][BioGrid]
86112 51242945 6801 / STRN 1891 13.75% 86.540 striatin; calmodulin binding protein [Homo sapiens] 65 29 49.60% [GI][Gene][BioGrid]

Selecting [BioGrid interactions not found here] opens a new window with the details of the "missed interactions", as shown below.

Note that the overlap is performed after data filtering is applied, thus care should be taken when analyzing apparent lack of overlap. The example bellow shows the effect of the application of a stringent filter on "missed interactions".

Missing interactions: STRINGENT FILTER

BioGrid interactions not found Show found hits

Gene ID	Gene Name	Links	BioGrid Type
Bait ID: 389 Bait Gene ID: 85369 Bait Gene Name: FAM40A [Gene][BioGrid]			
4163	MCC	[Gene][BioGrid]	☆
5515	PPP2CA	[Gene][BioGrid]	△
5516	PPP2CB	[Gene][BioGrid]	△
5518	PPP2R1A	[Gene][BioGrid]	△
10494	STK25	[Gene][BioGrid]	△
11235	PDCD10	[Gene][BioGrid]	△
29888	STRN4	[Gene][BioGrid]	△
57464	FAM40B	[Gene][BioGrid]	△
80342	TRAF3IP3	[Gene][BioGrid]	△

Missing interactions: NO FILTER

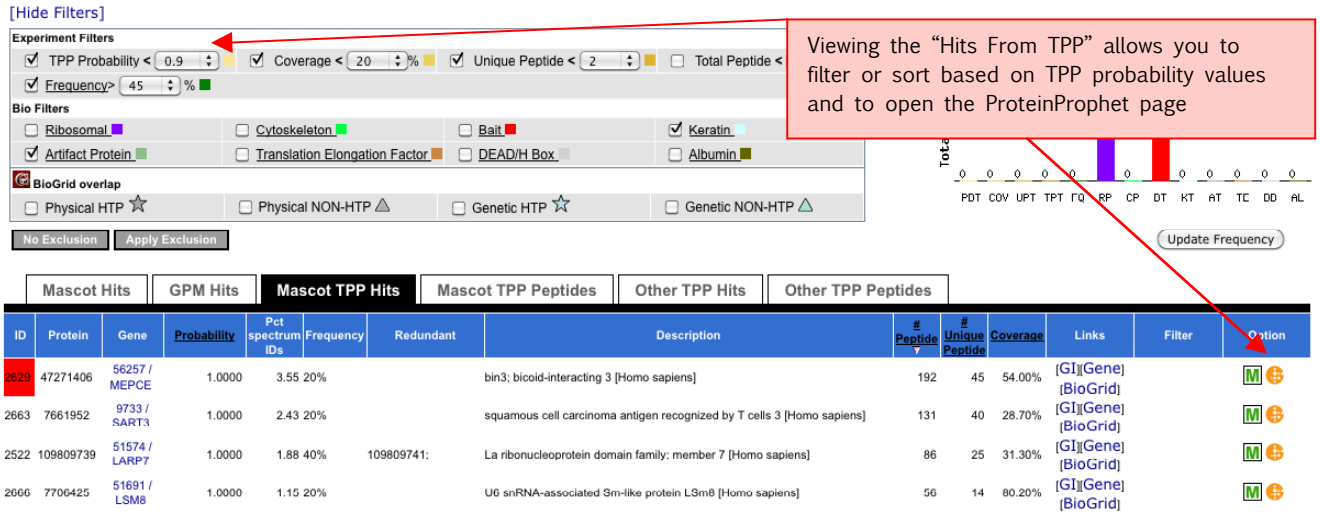
BioGrid interactions not found Show found hits

Gene ID	Gene Name	Links	BioGrid Type
Bait ID: 389 Bait Gene ID: 85369 Bait Gene Name: FAM40A [Gene][BioGrid]			
4163	MCC	[Gene][BioGrid]	☆
5516	PPP2CB	[Gene][BioGrid]	△
10494	STK25	[Gene][BioGrid]	△
80342	TRAF3IP3	[Gene][BioGrid]	△

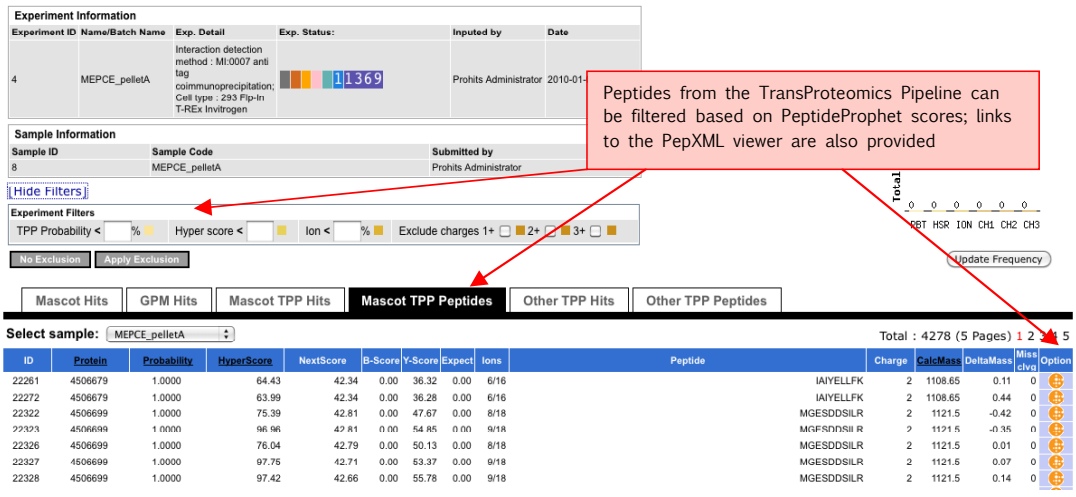
View and navigate hits from the TransProteomics Pipeline

The tabs located immediately above the search results table allow you to explore search results that have been parsed from the PeptideProphet and ProteinProphet components of the TPP.

In the page "Mascot TPP hits", different filtering options based on the number of unique or total peptides, as well as the probability values for the TPP have been implemented. A link to the TPP search result viewer is provided in the Option column of the table (orange Institute for Systems Biology icon): this opens up the standard ProteinProphet view, allowing further exploration of the data.


[Hide Filters]
Experiment Filters
TPP Probability < 0.9
Coverage < 20 %
Unique Peptide < 2
Total Peptide <
Frequency> 45 %
Bio Filters
Ribosomal
Cytoskeleton
Bait
Keratin
Artifact Protein
Translation Elongation Factor
DEAD/H Box
Albumin
BioGrid overlap
Physical HTP
Physical NON-HTP
Genetic HTP
Genetic NON-HTP
No Exclusion
Apply Exclusion
Viewing the "Hits From TPP" allows you to filter or sort based on TPP probability values and to open the ProteinProphet page
PDT COV UPT TPT FQ RP CP DT KT AT TE DD AL
Update Frequency
Mascot Hits
GPM Hits
Mascot TPP Hits
Mascot TPP Peptides
Other TPP Hits
Other TPP Peptides
ID Protein Gene Probability Pct spectrum IDs Frequency Redundant Description # Peptide # Unique Peptide Coverage Links Filter Option
2629 47271406 56257 / MEPCE 1.0000 3.55 20% bin3; bicoid-interacting 3 [Homo sapiens] 192 45 54.00% [GI][Gene] [BioGrid]
2663 7661952 9733 / SART3 1.0000 2.43 20% squamous cell carcinoma antigen recognized by T cells 3 [Homo sapiens] 131 40 28.70% [GI][Gene] [BioGrid]
2522 109809739 51574 / LARP7 1.0000 1.88 40% 109809741: La ribonucleoprotein domain family; member 7 [Homo sapiens] 86 25 31.30% [GI][Gene] [BioGrid]
2666 7706425 51691 / LSM8 1.0000 1.15 20% U6 snRNA-associated Sm-like protein LSm8 [Homo sapiens] 56 14 80.20% [GI][Gene] [BioGrid]

The "Mascot TPP Peptides" tab lists all of the parsed parameters at the peptide level, and provides some basic filtering options, as well as a link to the PepXML viewer.

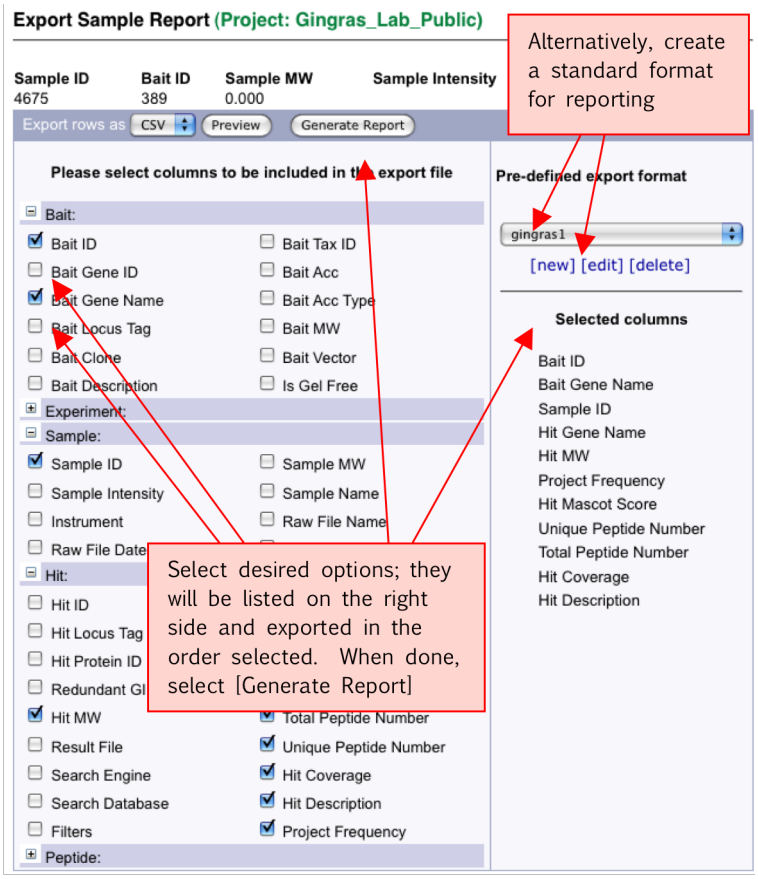

Experiment Information
Experiment ID Name/Batch Name Exp. Detail Exp. Status: Inputed by Date
4 MEPCE_pelletA Interaction detection method : MI:0007 anti tag coimmunoprecipitation; Cell type : 293 Flp-In T-REx Invitrogen Prohits Administrator 2010-01-
Sample Information
Sample ID Sample Code Submitted by
8 MEPCE_pelletA Prohits Administrator
[Hide Filters]
Experiment Filters
TPP Probability < % Hyper score < Ion < % Exclude charges 1+ 2+ 3+
No Exclusion
Apply Exclusion
Peptides from the TransProteomics Pipeline can be filtered based on PeptideProphet scores; links to the PepXML viewer are also provided
PBT HSR ION CH1 CH2 CH3
Update Frequency
Mascot Hits
GPM Hits
Mascot TPP Hits
Mascot TPP Peptides
Other TPP Hits
Other TPP Peptides
Select sample: MEPCE_pelletA
Total : 4278 (5 Pages) 1 2
ID Protein Probability HyperScore NextScore B-Score Y-Score Expect Ions Peptide Charge CalcMass DeltaMass Miss clvg Option
22261 4506679 1.0000 64.43 42.34 0.00 36.32 0.00 6/16 IAIYELLFK 2 1108.65 0.11 0
22272 4506679 1.0000 63.99 42.34 0.00 36.28 0.00 6/16 IAIYELLFK 2 1108.65 0.44 0
22322 4506699 1.0000 75.39 42.81 0.00 47.67 0.00 8/18 MGESDDSILR 2 1121.5 -0.42 0
22323 4506699 1.0000 96.96 42.81 0.00 54.85 0.00 9/18 MGESDDSILR 2 1121.5 -0.35 0
22326 4506699 1.0000 76.04 42.79 0.00 50.13 0.00 8/18 MGESDDSILR 2 1121.5 0.01 0
22327 4506699 1.0000 97.75 42.71 0.00 53.37 0.00 9/18 MGESDDSILR 2 1121.5 0.07 0
22328 4506699 1.0000 97.42 42.66 0.00 55.78 0.00 9/18 MGESDDSILR 2 1121.5 0.14 0

Viewing results using Cytoscape

At the top right corner on the Report page is a link to the molecular interaction visualization program Cytoscape. Clicking this link will upload the filtered data (with BioGrid interactions if this option is selected). Note that all mass spectrometry data will also be uploaded (you can use these parameters as attributes of the "edges" in Cytoscape). We will review Cytoscape requirements and basic information in the discussion of the "Comparison" function.

Export Sample report

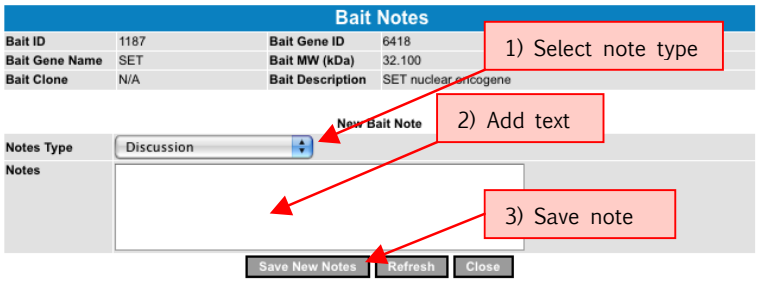
Selecting the "Export Sample Report" on the top right corner allows the user to export text (comma-separated values (CSV) or tab separated values (TSV)) files. Fields to be exported are user-defined and will be exported in the order selected. The user can also create pre-defined export formats that can be further modified. Note that this exports NON-FILTERED hits (filtered hits can be exported via the comparison tool).



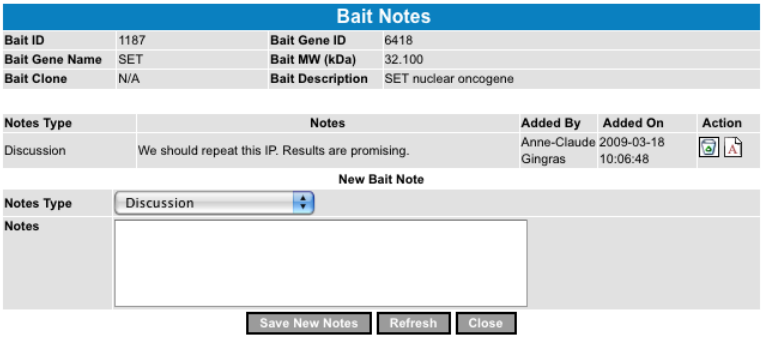
The exported file can be opened with Excel or similar software.

Using the Notes option

⇒ Click on the "callout" icon at the end of any bait row
This brings up the following window:



⇒ Add desired text, and press [Save New Notes]
The following screen can then be seen:



Only the person who entered the note is allowed to modify or delete it. Additional users can create additional comments on the same bait or sample.

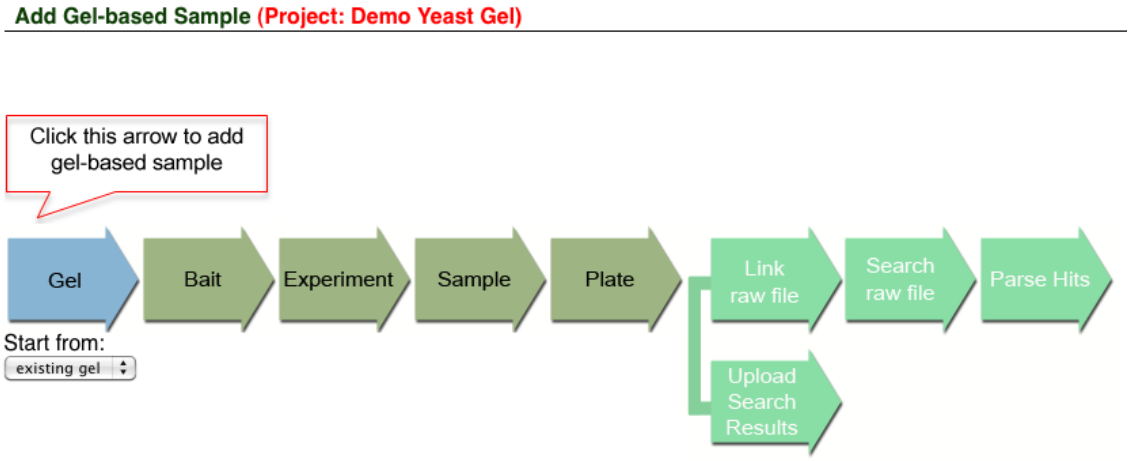
In addition to adding free text annotation (default "Discussion" note type), "Bait groups", "Experiment groups" or "Sample Groups" can be created for each project and are managed via the "Manage Protocols and Lists" option. Use the dropdown box to select the desired "Notes Types".

Creating gel-based samples

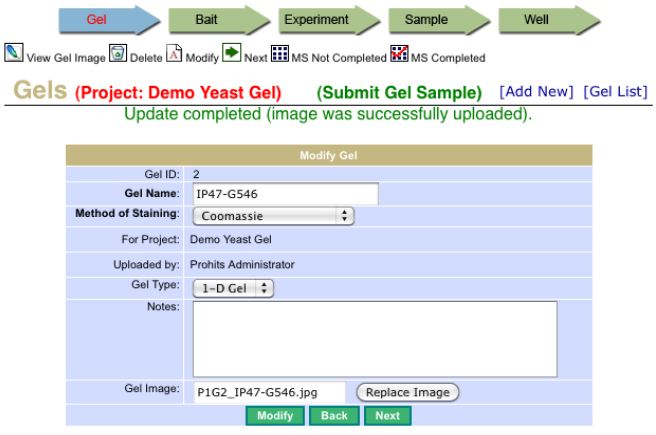
ProHits has functionality designed to track samples analyzed in a high-throughput manner from gel-based proteomics. Several of the steps are identical to the steps required to create samples for gel-free projects. Here we will briefly outline the major differences when entering gel-based samples. Note that you can add samples from in-gel digestion as "gel-free" – especially if you are only analyzing a few samples without the use of an autosampler.

Adding a "Gel-based" sample

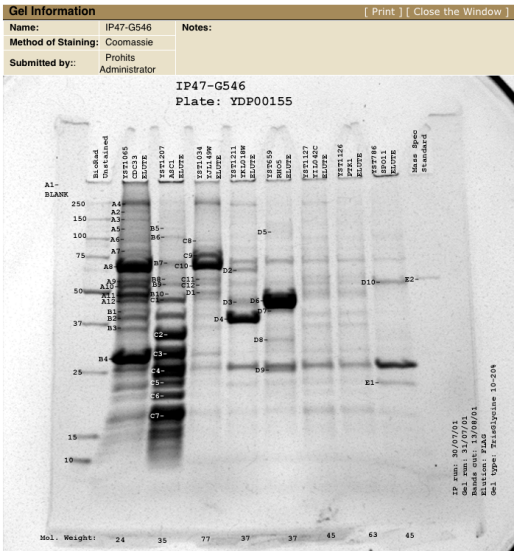
⇒ Select "Add Gel-based" sample from the left menu, and choose whether you will be starting from an existing gel, or create a new gel.



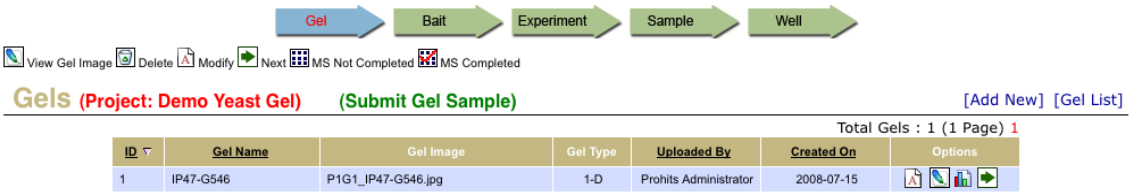
⇒ To create a new gel, add information required in bold, and upload the image of the gel.



While the image is not mandatory, it is highly recommended to link a well-annotated image of the gel.



After a gel is created, you can see the information via the "Report by Gel" function on the left menu.



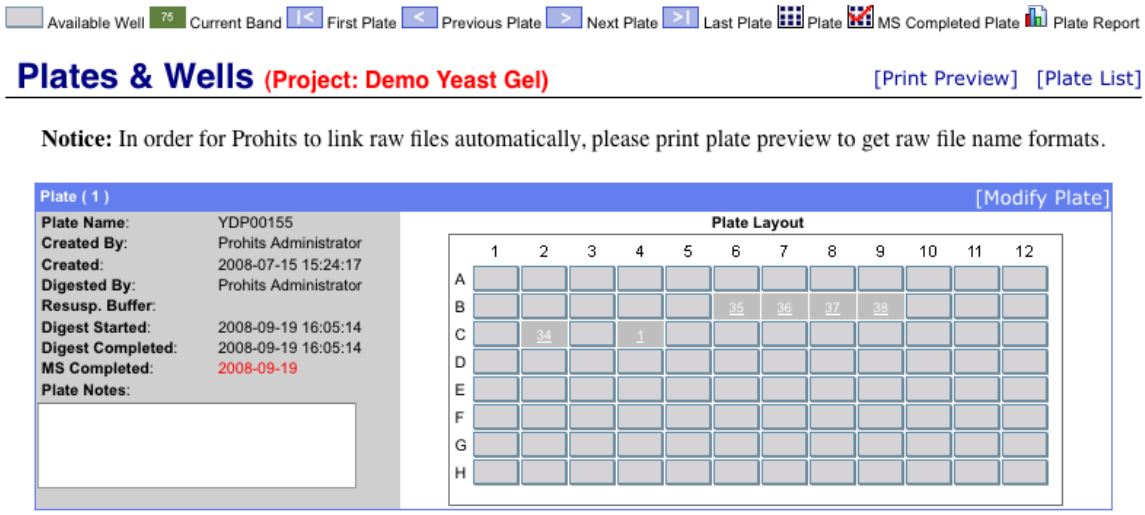
⇒ Use the green arrows in the "Options" field to enter baits from that gel (as shown in the gel-free sample section).

⇒ From each bait, define the Experimental Details, as shown in the gel-free section.

Clicking on the green arrow in the experimental details section will by default prompt you to define a lane on the gel, and guide you through the entry of individual band samples in the autosampler plate that you will use for data acquisition. Simply clicking on a plate well will create an associated sample – you can add the intensity of each band on the stained gel, as well as the approximate molecular weight.

Continue entering all desired bands from the selected lane, or use the navigation options at the top of the page to upload samples from the next lane, return to the list of all lanes, or return to the experimental description.

Opening the "Report by plate" and clicking the plate icon in the "Options" field, allows you to view your plate layout.



If you wish to use the "Auto-link" option to link your raw files from the Data Management module to the samples in Analyst, select [Print Preview].

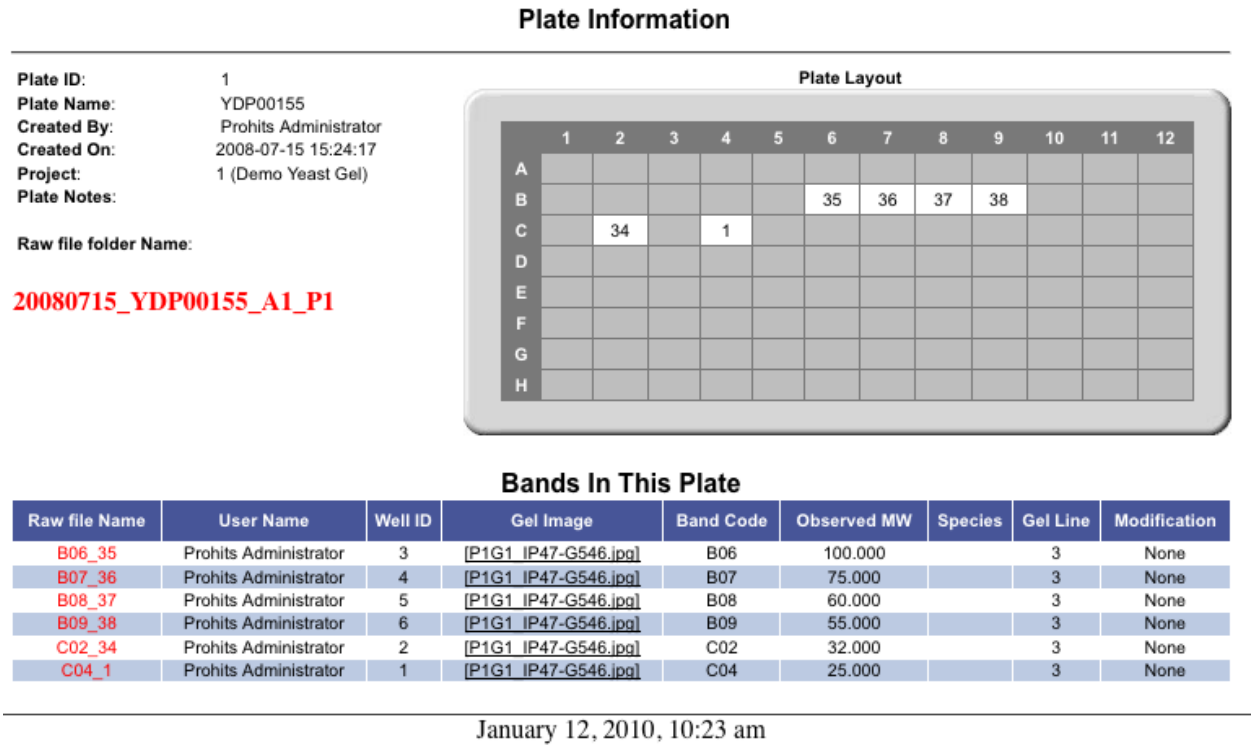
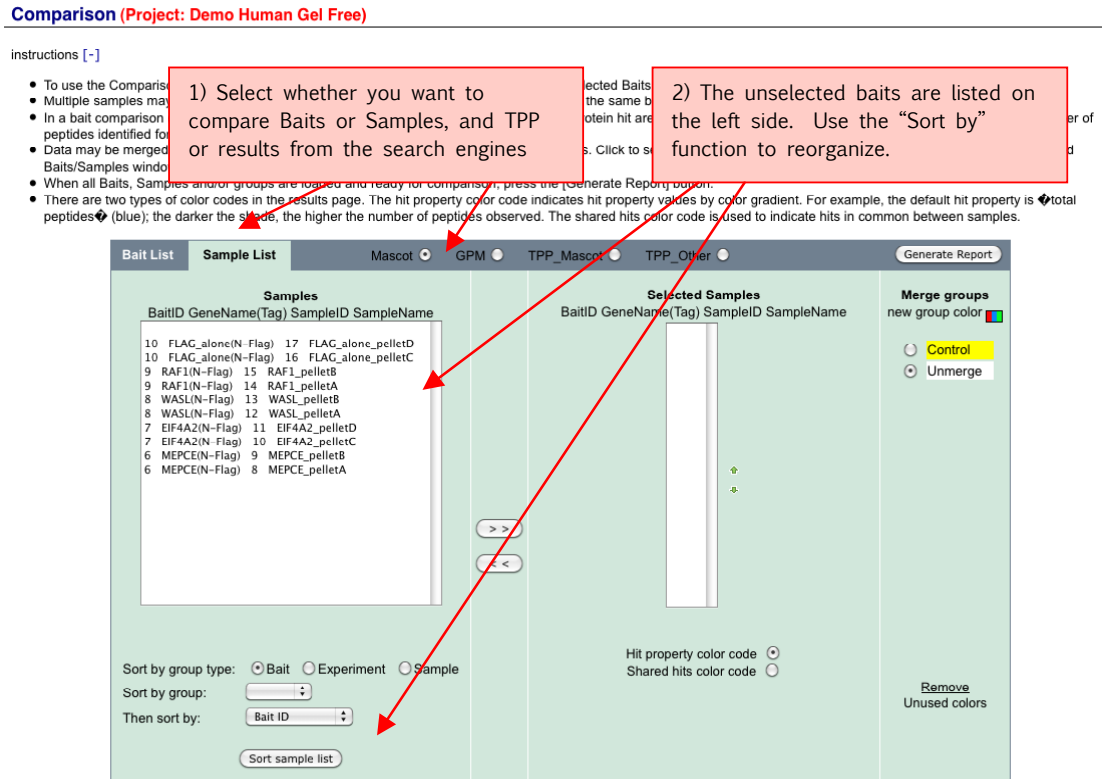


Plate Information
Plate ID: 1
Plate Name: YDP00155
Created By: Prohits Administrator
Created On: 2008-07-15 15:24:17
Project: 1 (Demo Yeast Gel)
Plate Notes:
Raw file folder Name:
20080715_YDP00155_A1_P1
Plate Layout
Bands In This Plate
Raw file Name | User Name | Well ID | Gel Image | Band Code | Observed MW | Species | Gel Line | Modification
B06_35 | Prohits Administrator | 3 | [P1G1_IP47-G546.jpg] | B06 | 100.000 | | 3 | None
B07_36 | Prohits Administrator | 4 | [P1G1_IP47-G546.jpg] | B07 | 75.000 | | 3 | None
B08_37 | Prohits Administrator | 5 | [P1G1_IP47-G546.jpg] | B08 | 60.000 | | 3 | None
B09_38 | Prohits Administrator | 6 | [P1G1_IP47-G546.jpg] | B09 | 55.000 | | 3 | None
C02_34 | Prohits Administrator | 2 | [P1G1_IP47-G546.jpg] | C02 | 32.000 | | 3 | None
C04_1 | Prohits Administrator | 1 | [P1G1_IP47-G546.jpg] | C04 | 25.000 | | 3 | None
January 12, 2010, 10:23 am

When setting up the acquisition on the mass spectrometer, the folder name (here **20080715_YDP00155_A1_P1**) as well as the Raw file names (e.g. **B06_35**) need to match these above.

The Comparison tool

ProHits has a built-in comparison tool that allows you to look at the results of several experiments side-by-side. You can perform comparisons at the bait level or at the sample level, and compare the results from the search engines (e.g. Mascot or X!Tandem) or the TPP. For this demonstration, we will perform a comparison at the sample level, using the Mascot search engine.



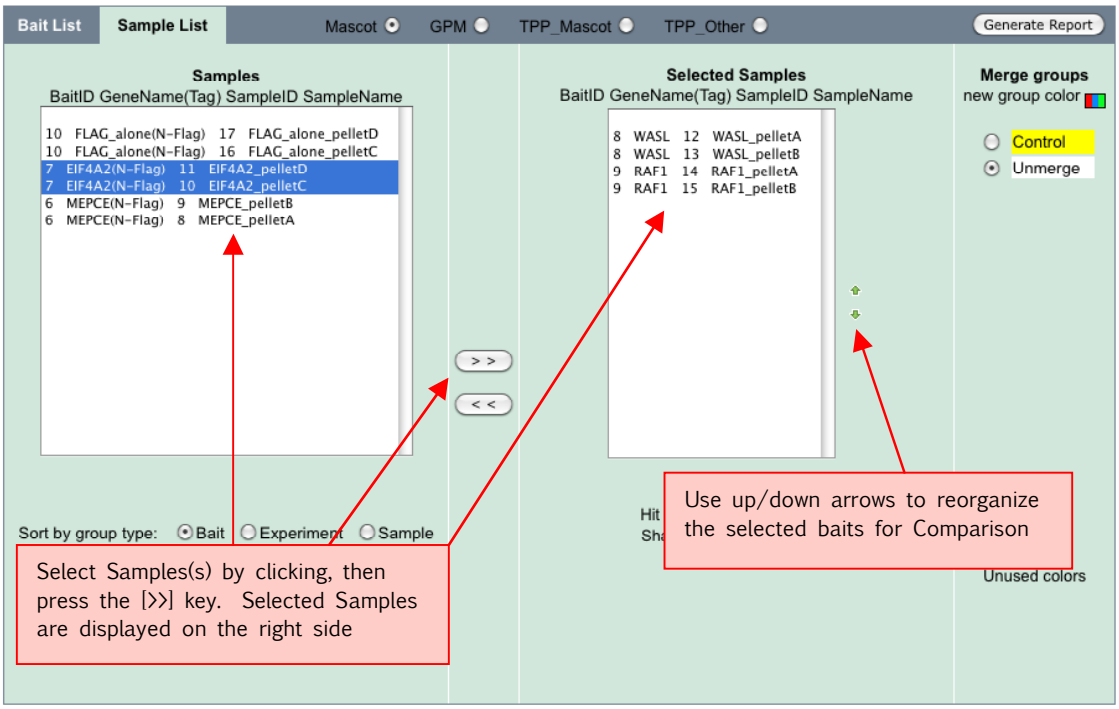
⇒ Select the desired baits to be compared

You can sort by Bait ID, Gene name, Protein ID, or by any of the user-defined flags that were used for the project.

⇒ Press the >> arrow button to transfer the baits to the "Selected Baits" window

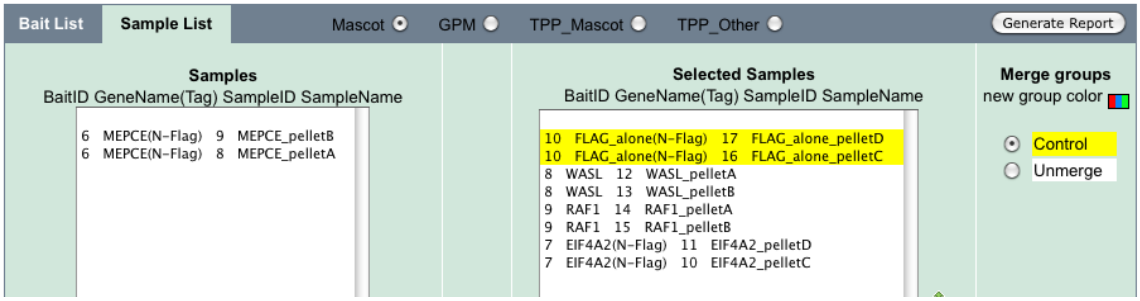
You can transfer files one at the time, or by large groups. The files are added to the list in the order selected. This will also be the order of the columns in the Comparison View.

Use the green up/down arrows on the right hand side to reorganize the sort order. Individual Baits or Groups of Baits can be reorganized.



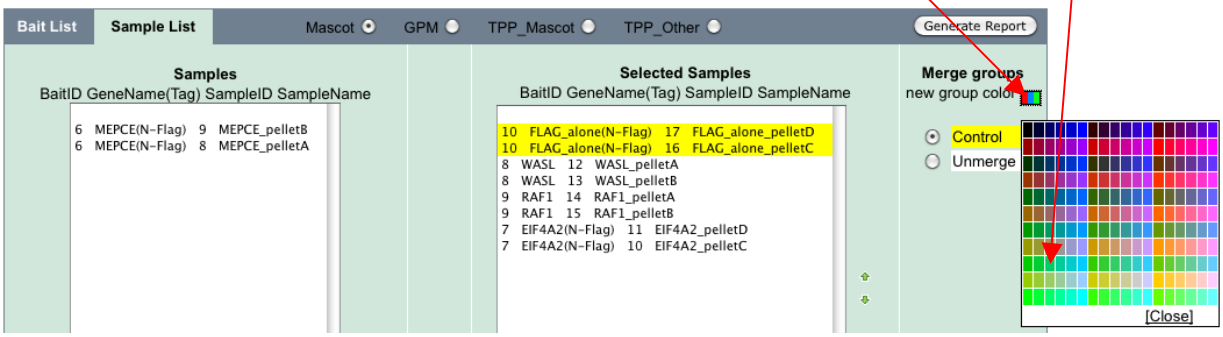
Merging files prior to Comparison

Additional options are available that provide merging options for two or more files. *Please note that the merging is a very simple process that simply reports the best hits for the item but does not do any recalculation. If the selected display option in the report is the Mascot score, the best scoring hit will be listed; if the selected display option is based on spectral counts, the hit with the highest spectral counts will be reported*. The merging function allows you to group two or more control runs (click on the "Control" button before transferring the selected files).

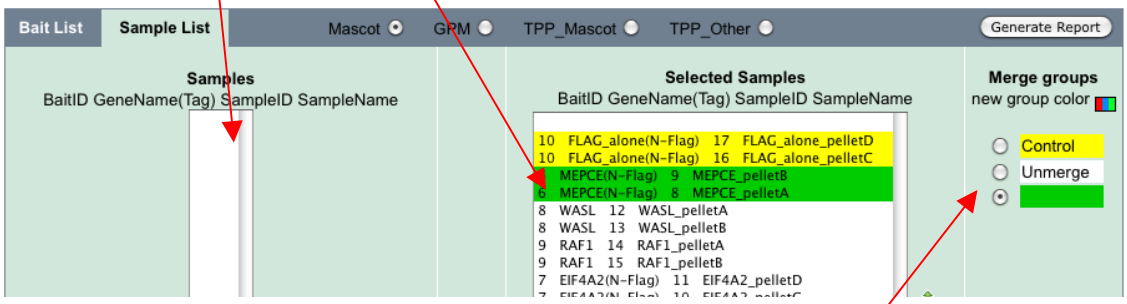


You can group any set of additional files by first clicking on the multicolour icon to select a new group, then transferring the given files to the right side. The listing order will be as follows: The control group will be listed first, followed by all other groups in the order selected by the user, followed by all individual entries in the order selected by the user. Note that within the same group, hits will be combined, and only the maximal value for each of the properties will be reported.

1) To combine multiple Baits into one group, click first on the multicolour icon and select desired colour



2) Select and transfer the files as above; the selected files are now grouped and highlighted with the desired colour

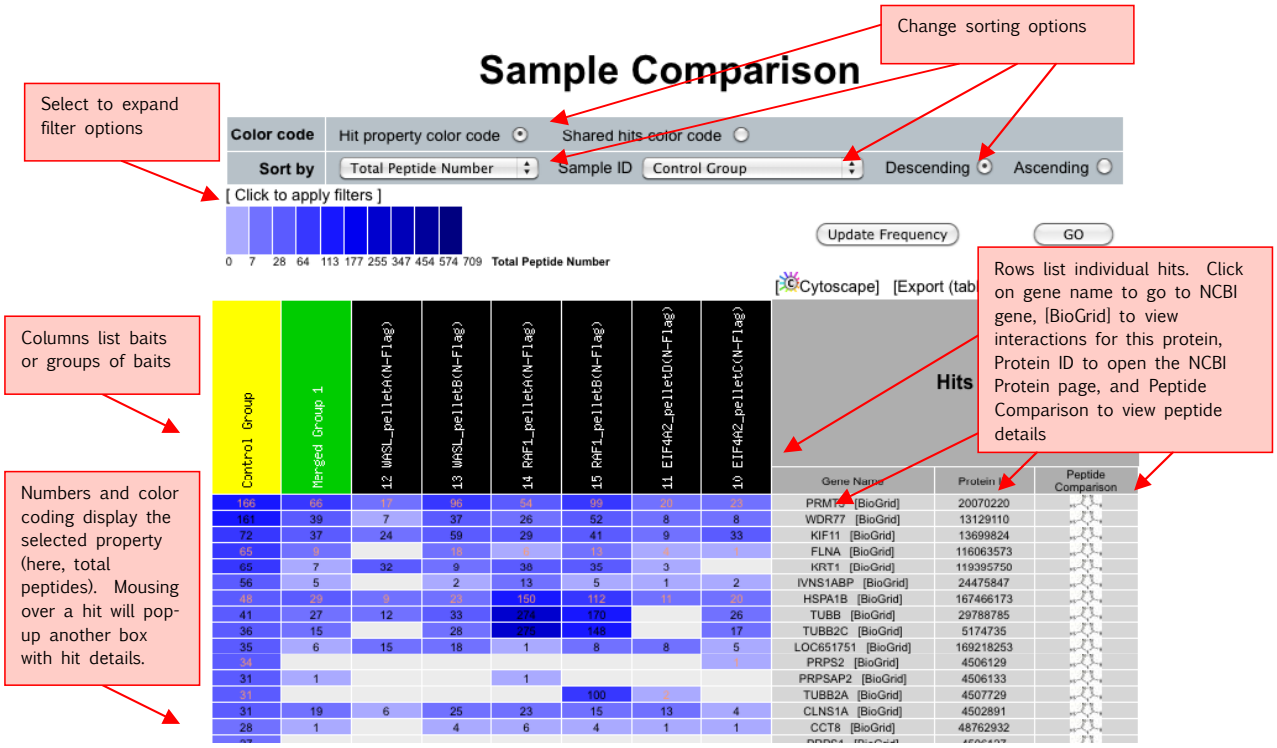


3) Multiple groups of this type can be created in the same manner. If you want to add single files, first select the "unmerge" button, then browse in the file list

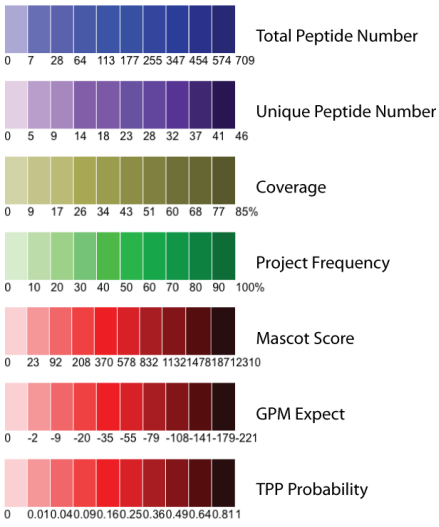
⇒ When you are done adding all desired baits and/or bait groups, press [Generate Report]
This will open a new window, the Comparison page.

Comparison page

When you open the Bait Comparison page, you will see an unfiltered view of the hits. Each column represents a different sample or bait (or group of samples or baits if the "merge" function was used). The rows represent each of the hits detected across the *n* samples or baits. Clicking on the Gene Name will take you to NCBI Gene; selecting [BioGrid] will open the BioGrid entry for the given protein; clicking on the number in the Protein ID field will bring you to the Entrez Protein page. The last column allows you to compare the peptides identified across the bait purifications.


Change sorting options
Sample Comparison
Select to expand filter options
Color code
Hit property color code
Shared hits color code
Sort by
Total Peptide Number
Sample ID
Control Group
Descending
Ascending
[Click to apply filters]
0 7 28 64 113 177 255 347 454 574 709 Total Peptide Number
Update Frequency
GO
[Cytoscape] [Export (tab
Hits
Rows list individual hits. Click on gene name to go to NCBI gene, [BioGrid] to view interactions for this protein, Protein ID to open the NCBI Protein page, and Peptide Comparison to view peptide details
Columns list baits or groups of baits
Control Group
Merged Group 1
12 WASL_pelletA(N-Flag)
13 WASL_pelletB(N-Flag)
14 RAF1_pelletA(N-Flag)
15 RAF1_pelletB(N-Flag)
11 EIF4H2_pelletD(N-Flag)
10 EIF4H2_pelletC(N-Flag)
Gene Name
Protein ID
Peptide Comparison
Numbers and color coding display the selected property (here, total peptides). Mousing over a hit will pop-up another box with hit details.
PRMT5 [BioGrid] 20070220
WDR77 [BioGrid] 13129110
KIF11 [BioGrid] 13699824
FLNA [BioGrid] 116063573
KRT1 [BioGrid] 119395750
IVNS1ABP [BioGrid] 24475847
HSPA1B [BioGrid] 167466173
TUBB [BioGrid] 29788785
TUBB2C [BioGrid] 5174735
LOC651751 [BioGrid] 169218253
PRPS2 [BioGrid] 4506129
PRPSAP2 [BioGrid] 4506133
TUBB2A [BioGrid] 4507729
CLNS1A [BioGrid] 4502891
CCT8 [BioGrid] 48762932

The default display is with Total Peptide Numbers (spectral counts), and the default sorting option is by descending number of spectra, starting by the left-most bait or group. Note that these sorting options can be modified. In particular, ProHits recalculates and sorts using the following parameters:



In addition to the sorting options, ProHits Comparison allows you to filter your data in a manner similar to the filtering options in the Bait Report page.

⇒ To access the filtering option, select [Click to apply filters].

An expanded menu allows you to select criteria for removal of proteins from the Comparison list.

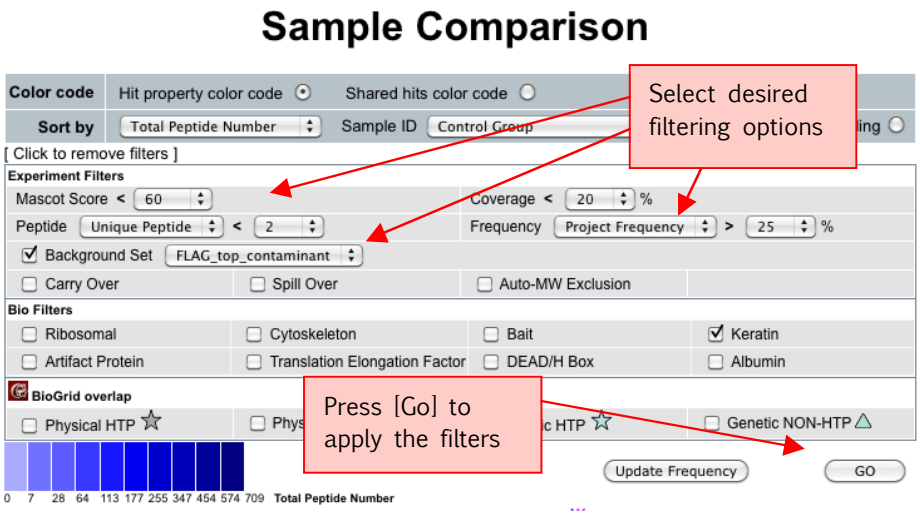
⇒ Select desired parameters

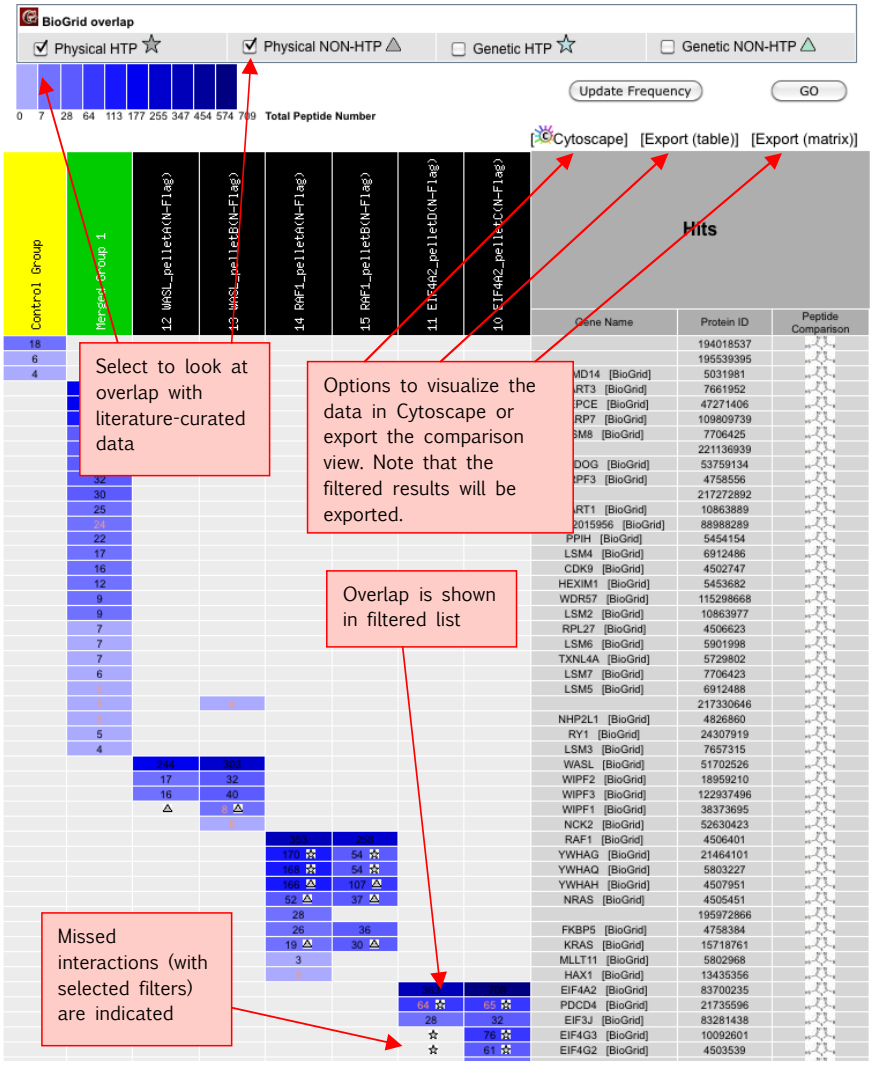
⇒ Select to highlight the BioGrid overlap if desired

⇒ To apply filters, press [Go]

This generates a modified list, similar to the process described in the Bait report section. If selected, the overlap with BioGrid is indicated by stars or triangles in the list below.

Note that mousing over any of the entries shown below will pop up a menu box listing the scoring details.

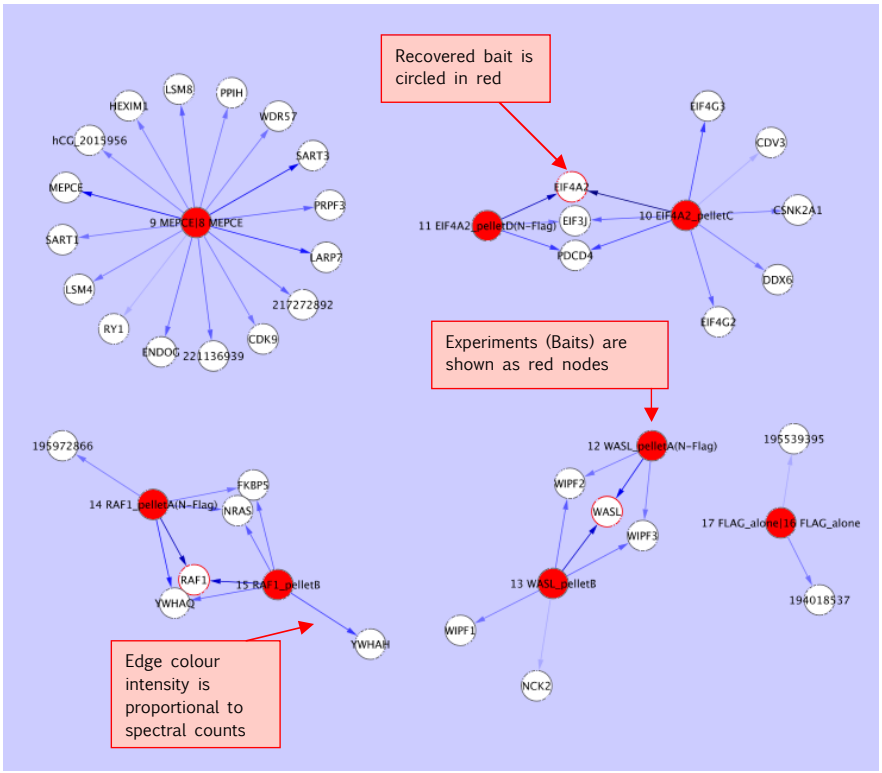




Using Cytoscape directly from ProHits comparison

ProHits allows you to visualize your data using Cytoscape. If using the ProHits filters, the data post-filtering will be displayed (changing the filter will modify the display). If the BioGrid overlap function has been selected, the resulting Cytoscape view will incorporate both your mass spectrometry data, the overlap between your mass spectrometry data and data in BioGrid, and data detected only in BioGrid (including interactions amongst first neighbours of the hits). The colour-coding (see below) allows you to identify the source of the data.

Before you can use the Cytoscape plug-in, you need to have the Runtime Environment (JRE) installed on your local computer (you can use the following URL to test whether your computer has a functional JRE: http:///www.java.com/en/download/help/testum.xml). The first time that you click the "Cytoscape" icon, Cytoscape will be installed on your local computer. Press the [Cytoscape] link immediately above the table to open the current interaction file in Cytoscape. The baits are indicated by red nodes (alongside the unique bait identifier), and the recovery of baits in a purification is indicated by circling the white baits in red. The colour-coding of the arrows is mapped to the spectral counts, as shown above, and all peptide annotation is encoded as an edge attribute. Note that if the "Overlap with BioGrid" function has been selected, interactions specific to your dataset will be still shown in blue, interactions that overlap between your dataset and BioGrid will be shown in green, while BioGrid-only interactions will be displayed in white.



The original image is a circular layout; in the example shown here, this has simply been converted to a spring-embedded layout, with weight on the edge (unique peptide).

Note that all of the standard Cytoscape tools are available.

Other export options

You may also wish to launch Cytoscape (or additional network viewers) from an Excel Table, in which you can add annotation or other mapping options. To do so, use the **[Export (table)]** option, also located at the top of the table. This will create a .csv file that can be opened and modified in Excel. The file will be displayed as a bait>hit list with each subsequent column listing a separate parameter. These lists are easily opened using a stand-alone Cytoscape version.

	A	B	C	D	E	F	G	H	I	J	K
1	Sample ID	Bait Gene Name	Hit Gene Name	Hit Gene ID	Hit Protein ID	Hit Score	Peptide Number	Unique Peptide Number	Coverage	Frequency	Shared Frequency
2	10	EIF4A2_pelletC(N-Flag)	EIF4A2	1974	83700235	1759	709	24	71.5	60	20
3	11	EIF4A2_pelletD(N-Flag)	EIF4A2	1974	83700235	1720	363	25	64.6	60	20
4	14	RAF1_pelletA(N-Flag)	RAF1	5894	4506401	1620	353	20	50.9	20	20
5	13	WASL_pelletB(N-Flag)	WASL	8976	51702526	1320	303	21	60.4	20	20
6	15	RAF1_pelletB(N-Flag)	RAF1	5894	4506401	1560	258	20	53.9	20	20
7	12	WASL_pelletA(N-Flag)	WASL	8976	51702526	1457	244	23	57.2	20	20
8	9	MEPCE_pelletB(N-Flag)	SART3	9733	7661952	1683	242	24	35.8	20	20
9	9	MEPCE_pelletB(N-Flag)	MEPCE	56257	47271406	1480	184	24	53.7	20	20
10	9	MEPCE_pelletB(N-Flag)	LARP7	51574	109809739	1121	176	19	33.5	20	20
11	14	RAF1_pelletA(N-Flag)	YWHAG	7532	21464101	321	170	3	27.1	20	20
12	14	RAF1_pelletA(N-Flag)	YWHAQ	10971	5803227	455	168	4	26.5	20	20
13	14	RAF1_pelletA(N-Flag)	YWHAH	7533	4507951	295	166	3	22.8	20	20
14	8	MEPCE_pelletA(N-Flag)	MEPCE	56257	47271406	1106	152	19	43.8	20	20
15	15	RAF1_pelletB(N-Flag)	YWHAH	7533	4507951	472	107	4	33.3	20	20
16	8	MEPCE_pelletA(N-Flag)	SART3	9733	7661952	1463	105	20	26.9	20	20
17	10	EIF4A2_pelletC(N-Flag)	EIF4G3	8672	10092601	1598	76	25	23.3	20	10

[Export (matrix)] provides a view similar to that displayed in the Comparison page, with the option to export only the parameter currently displayed (e.g. spectral counts), or the option to list all parameters inside each cell. Again, a .csv file that can be opened and modified in Excel will be created.

View only the displayed value (here = total peptide counts):

	A	B	F	G	H
1	Generated date: 2010-January-15		Project Name: Demo Human Gel Free		
2	PID:SC(PT-PU-C%-F%-SF%)		Peptide ID:Score(Total Peptide Number-Unique Peptide Numbe-Cover		
3					
4			17 FLAG_alone_pelletD	8 MEPCE_pelletA(N-Fl	9 MEPCE_pelletB(N-Flag
5					
6	Gene ID	Gene Name	Total Peptide Number	Total Peptide Number	Total Peptide Number
7	194018537		18		
8	195539395		6		
9	10213	PSMD14	4		
10	56257	MEPCE		152	184
11	9733	SART3		105	242
12	51574	LARP7		53	176
13	51691	LSM8		49	35
14	217272892			26	30
15	9129	PRPF3		26	32
16	221136939			25	40
17	25804	LSM4		17	11
18	1025	CDK9		16	
19	2021	ENDOG		15	34
20	10614	HEXIM1		12	
21	57819	LSM2		8	9

View all parameters:

	A	B	D	E	F	G	H
1	Generated date: 2010-January-15		Project Name: Demo Human Gel Free				
2	PID:SC(PT-PU-C%-F%-SF%)		Peptide ID:Score(Total Peptide Number-Unique Peptide Numbe-Coverage-Frequency-Sub Frequency)				
3							
4				16 FLAG_alone_pelletC(N-Flag)	17 FLAG_alone_pelletD(N-Flag)	8 MEPCE_pelletA(N-Flag)	9 MEPCE_pelletB(N-Flag)
5							
6	Gene ID	Gene Name	Protein ID	Total Peptide Number	Total Peptide Number	Total Peptide Number	Total Peptide Number
7	194018537		194018537		18		
8	195539395		195539395		6		
9	10213	PSMD14	5031981		4		
10	56257	MEPCE	47271406			152	184
11	9733	SART3	7661952			105	242
12	51574	LARP7	109809739			53	176
13	51691	LSM8	7706425			49	35
14	217272892		217272892			26	30
15	9129	PRPF3	4758556			26	32
16	221136939		221136939			25	40
17	25804	LSM4	6912486			17	11
18	1025	CDK9	4502747			16	
19	2021	ENDOG	53759134			15	34
20	10614	HEXIM1	5453682			12	
21	57819	LSM2	10863977			8	9

Zoom of the details inside each cell:

PID:SC(PT-PU-C%-F%-SF%)

56790935:500(16-10-20.70-11.88-75)

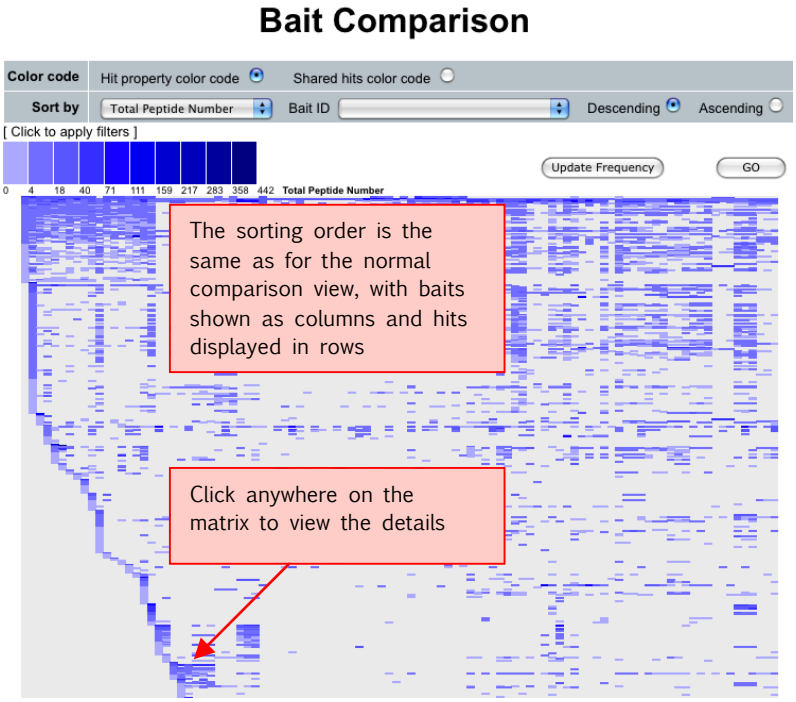
Legend:
PID: Protein ID (NCBI Entrez Protein)
SC: Mascot Score
PT: Total number of peptides
PU: Number of Unique peptides
C%: Percentage of the protein sequenced
F%: Frequency of occurrence of the protein in the entire dataset
SF%: Frequency of occurrence of the protein amongst compared baits/samples

Comparing larger numbers of baits

ProHits also allows you to visualize larger numbers of experiments.

⇒ Select the baits or samples to be compared and press [Generate Report]

A heat-map view of the data will be generated.


Bait Comparison
Color code
Hit property color code
Shared hits color code
Sort by
Total Peptide Number
Bait ID
Descending
Ascending
[Click to apply filters]
0 4 18 40 71 111 159 217 283 358 442 Total Peptide Number
Update Frequency
GO
The sorting order is the same as for the normal comparison view, with baits shown as columns and hits displayed in rows
Click anywhere on the matrix to view the details

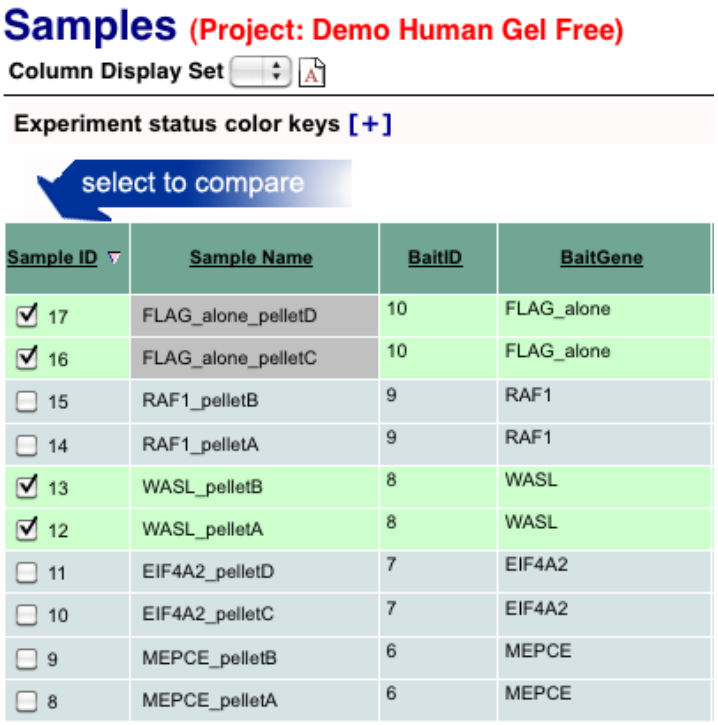
⇒ Click anywhere on the map to expand and view names and other details

Note, however, that due to file size, the [Cytoscape] option is not available with this heat map view. The [Export(table)] option is still available, however, and can allow you to upload data into a stand-alone Cytoscape session (the [Export(matrix)] function is also available). Note that due to large file sizes, these export functions may run slowly.

For additional export functionalities, you can go back to the main Analyst module, and select the "Export Hits" option from the left-hand menu.

Automatically adding baits for comparison from the baits or sample report list pages

ProHits allows you to select baits or samples to be added to the comparison page while working on other pages. To use this option, simply click the box located to the left side of each sample in the sample list or by the bait in the bait list.



Sample ID	Sample Name	BaitID	BaitGene
☑ 17	FLAG_alone_pelletD	10	FLAG_alone
☑ 16	FLAG_alone_pelletC	10	FLAG_alone
☐ 15	RAF1_pelletB	9	RAF1
☐ 14	RAF1_pelletA	9	RAF1
☑ 13	WASL_pelletB	8	WASL
☑ 12	WASL_pelletA	8	WASL
☐ 11	EIF4A2_pelletD	7	EIF4A2
☐ 10	EIF4A2_pelletC	7	EIF4A2
☐ 9	MEPCE_pelletB	6	MEPCE
☐ 8	MEPCE_pelletA	6	MEPCE

The selected sample (or baits) will be automatically added to the "Selected Samples" and "Selected Baits" pages of the Comparison view. Note that if a bait is selected, all samples corresponding to this bait will automatically be added to the comparison view.

You can keep browsing and adding baits or samples for Comparison as you go. These will stay selected for the duration of your session, or until you manually remove them from the Comparison page.

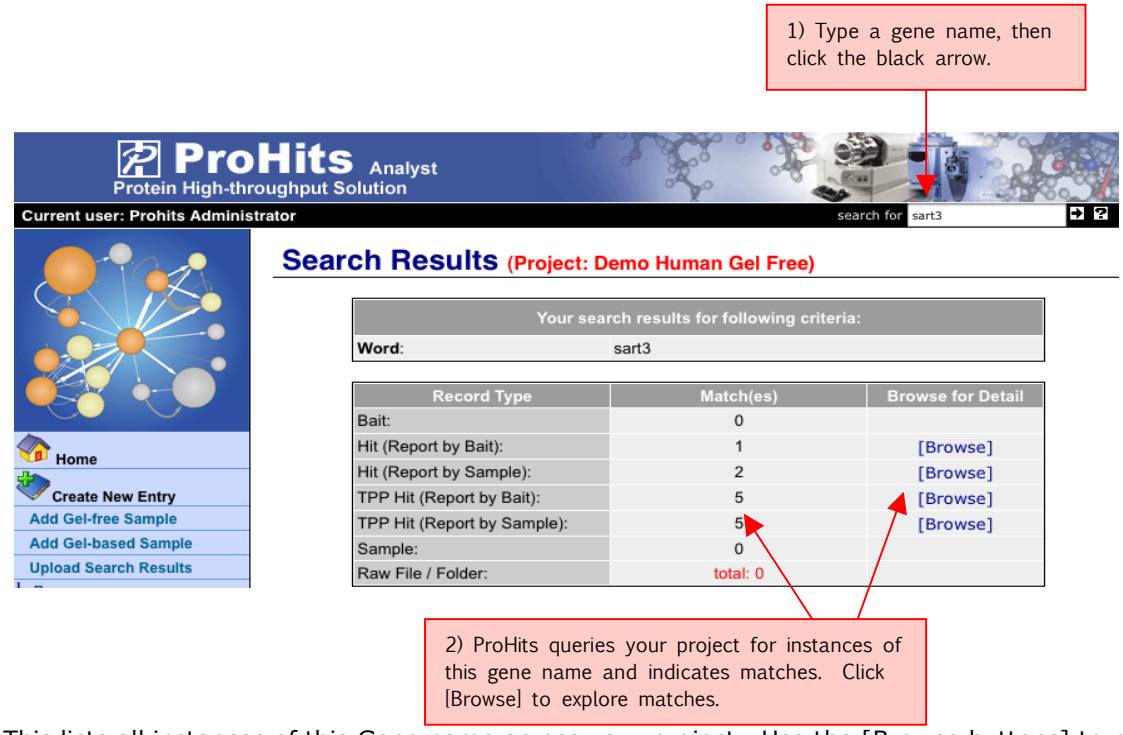
Search options

ProHits Analyst allows you to perform simple searches (for individual Gene Names) or Advanced searches (for multiple gene names or keywords in the protein description field or controlled vocabulary). Here, we will briefly review these options:

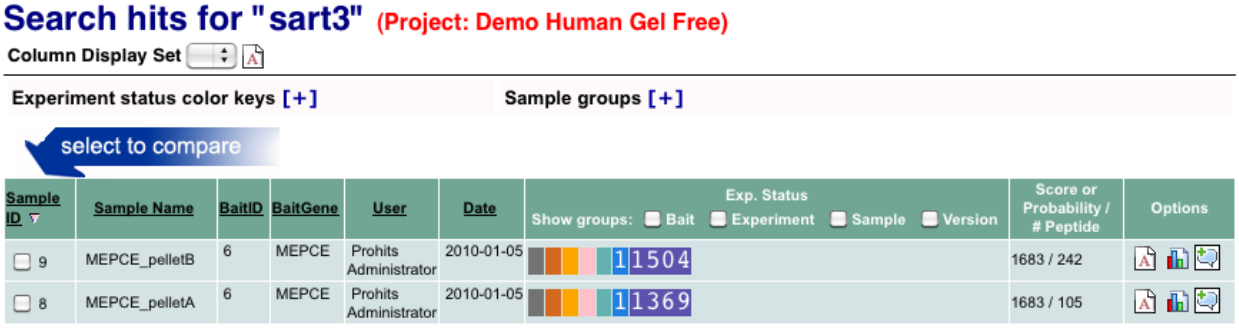
Simple Search (Gene name)

ProHits has a simple search function that is located at the upper corner of the Analyst module main page.

⇒ Enter an official Gene Name, then press the right pointing arrow.


1) Type a gene name, then click the black arrow.
ProHits Analyst
Protein High-throughput Solution
Current user: Prohits Administrator
search for sart3
Home
Create New Entry
Add Gel-free Sample
Add Gel-based Sample
Upload Search Results
Search Results (Project: Demo Human Gel Free)
Your search results for following criteria:
Word: sart3
Record Type | Match(es) | Browse for Detail
Bait: 0
Hit (Report by Bait): 1 [Browse]
Hit (Report by Sample): 2 [Browse]
TPP Hit (Report by Bait): 5 [Browse]
TPP Hit (Report by Sample): 5 [Browse]
Sample: 0
Raw File / Folder: total: 0
2) ProHits queries your project for instances of this gene name and indicates matches. Click [Browse] to explore matches.

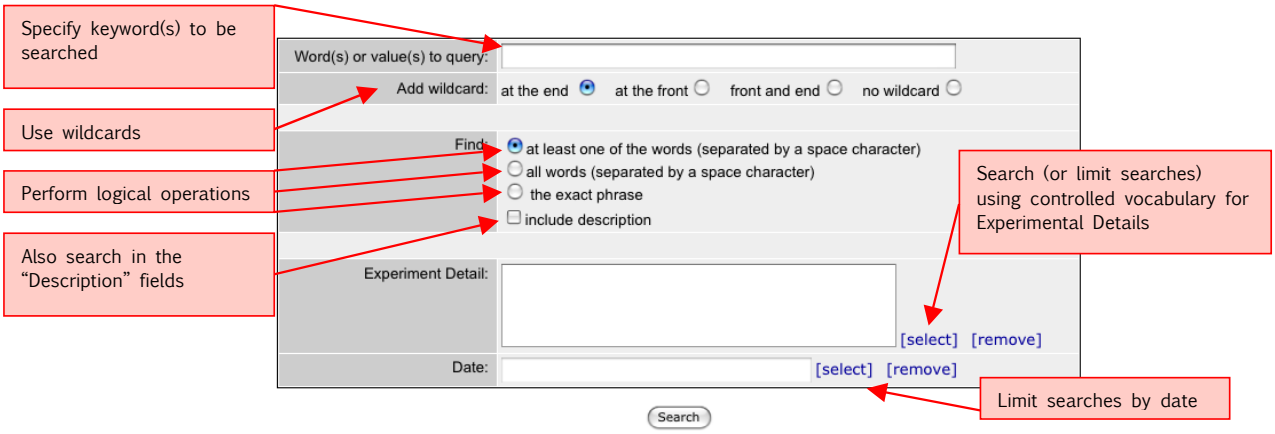
This lists all instances of this Gene name across your project. Use the [Browse buttons] to navigate through the data. Below, we have expanded the "Hit (Report by Sample)" option. The gene SART3 was identified in both of the MEPCE biological replicates. Note that the column "Score of Probability/ # Peptides" refers to the score from the search engines (or TPP) and the total number of peptides identified for SART3 in the MEPCE runs.


Search hits for "sart3" (Project: Demo Human Gel Free)
Column Display Set
Experiment status color keys [+]
Sample groups [+]
select to compare
Sample ID | Sample Name | BaitID | BaitGene | User | Date | Exp. Status Show groups: Bait Experiment Sample Version | Score or Probability / # Peptide | Options
9 MEPCE_pelletB 6 MEPCE Prohits Administrator 2010-01-05 11504 1683 / 242
8 MEPCE_pelletA 6 MEPCE Prohits Administrator 2010-01-05 11369 1683 / 105

Advanced Search

The Advanced Search function can be accessed from the menu bar. This function allows you to search for keywords (or combinations of keywords) and retrieve entries across the following categories: **Baits**, **Hits**, **Samples**, **Gels**, **Raw Files** and **Auto Search**.

In the simplest sense, you can use the Advanced search in a manner similar to the Simple search, i.e. to retrieve entries associated with a gene name. You can use "wildcards", either at the front, at the end, or both at the front and end of your query. Note that using wildcards (especially at the front) decreases search speed.

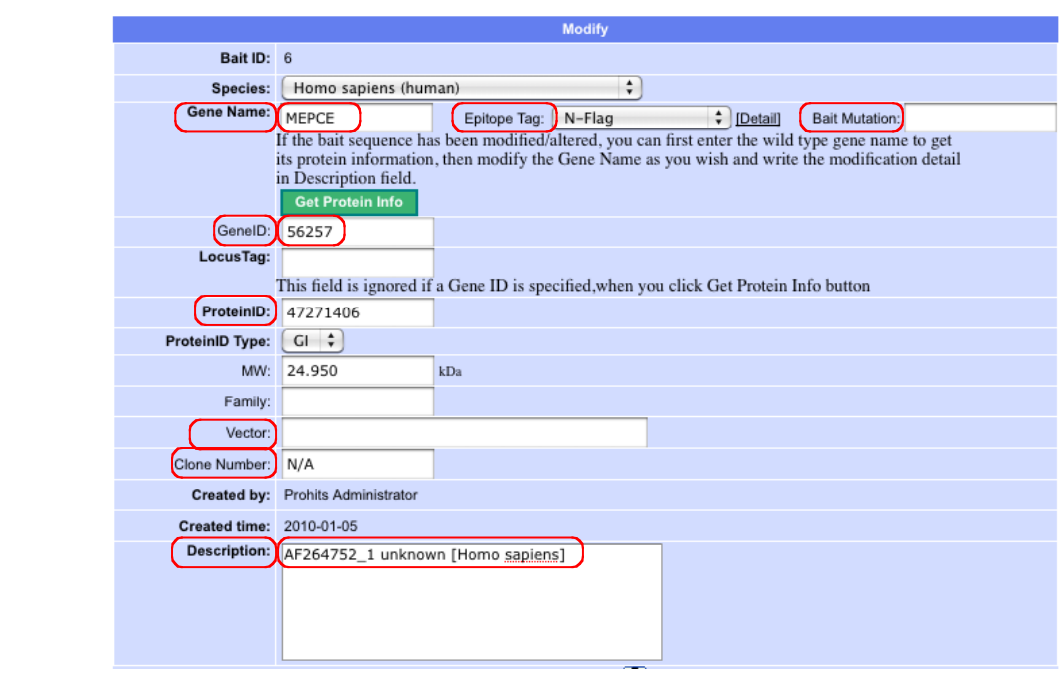


This will return a list of results that you can then explore further by selecting the [Browse] option for each of the categories, as for the simple search.

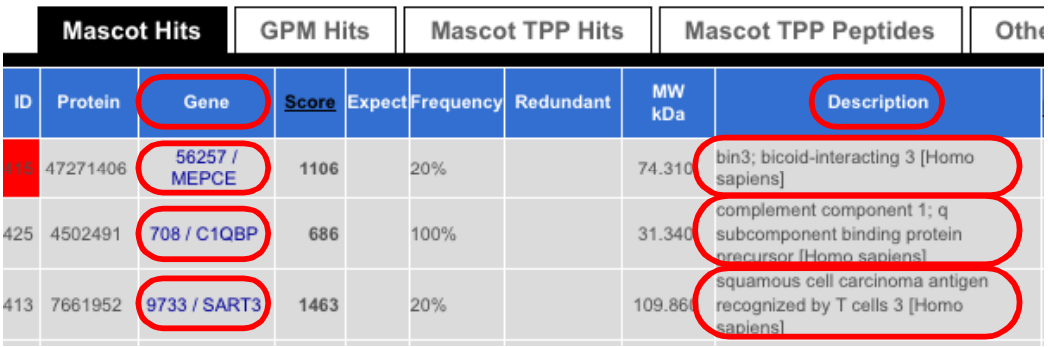
Other keywords that can be searched:

In addition to the Gene Name, different keywords can be searched. The fields searched depend upon the category, as defined below:

1) **Bait** (the keywords were detected in the entry for a bait – fields searched are "Gene Name", "Gene ID", "Locus Tag", "Protein ID", "Epitope Tag", "Bait Mutation", "Clone Number", "Vector", with optionally, Bait "Description"). The searched fields are indicated by red ovals below:


Baits (Project: Demo Human Gel Free)
Modify
Bait ID: 6
Species: Homo sapiens (human)
Gene Name: MEPCE
Epitope Tag: N-Flag
[Detail]
Bait Mutation:
If the bait sequence has been modified/altered, you can first enter the wild type gene name to get its protein information, then modify the Gene Name as you wish and write the modification detail in Description field.
Get Protein Info
GeneID: 56257
LocusTag:
This field is ignored if a Gene ID is specified,when you click Get Protein Info button
ProteinID: 47271406
ProteinID Type: GI
MW: 24.950 kDa
Family:
Vector:
Clone Number: N/A
Created by: Prohits Administrator
Created time: 2010-01-05
Description: AF264752_1 unknown [Homo sapiens]

2) **Hits** (the keywords were detected in the hits list – field searched is "Gene" Name, with, optionally, Protein "Description"). *You can similarly see the hits across TPP results.* The searched fields are indicated by red ovals below.



3) **Sample** (the keywords were detected in the user-defined "Sample Name")
4) **Gel** (the keywords were detected in the fields "Gene Name", "Gene Image", and "Lane Code")
5) **Raw files** (the keywords were detected in "File Name" or "Folder Name"). This brings you to the "Data management" module, and lists the folders / files bearing the selected keywords.
6) **Auto Search** (the keywords were detected in "Search Task Name"). This brings you to the "Data management" module, and lists the search tasks bearing the selected keywords.

Searching Bait/Protein Description: You can search for a keyword inside the Description field (e.g. "squamous" in the example above), by allowing wildcards on both sides. In other words, the entire field is captured (not individual words), and any partial field (e.g. "squamous" or "carcinoma") must be preceded and/or followed by wildcards. Note again that such searches may be very slow.

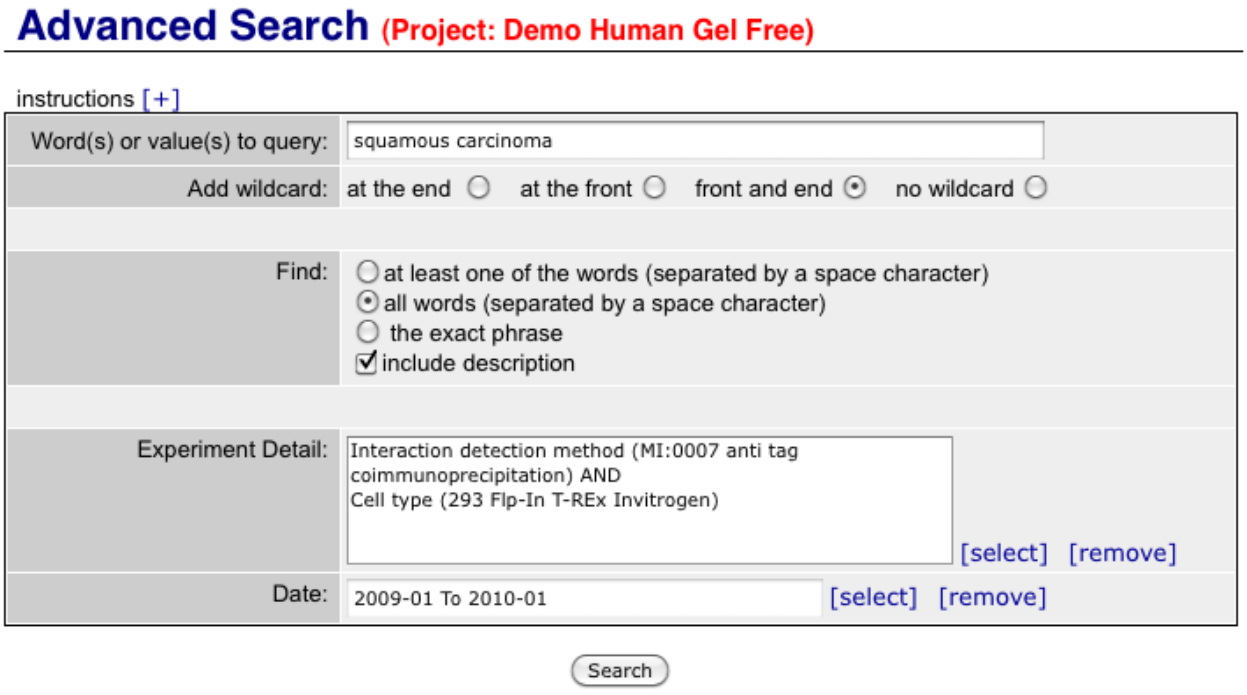
Searching in Experimental Details (controlled vocabularies): The search function also allows you to search (or limit your searches) based on selected controlled vocabulary. Simply press [Select] (bottom right corner of the Experimental Detail section). This will take you to the Experimental Details/controlled vocabulary section where you can select categories/values to be passed to the Advanced search page.

Restricting searches by date: You can restrict search results by date. Simply press the [select] button in the Date field to open a drop-menu.

Using logical operations: You can combine several keywords (simply separate them by spaces), to search for "at least one of the words", "all words" (in any order), or "the exact phrase" within a field, such as "Description". Note that the "all words" and "exact phrase" operations only apply within a field. Alternatively, you can use the "at least one of the words" option to search for different keywords even across different fields. This will generate a list of results that will be the union of the separate lists.

Hits searches returning too many results: Note that there is a limit of 3000 to search results. Try narrowing down your search parameters and try again.

Example: Searching for squamous AND carcinoma in 293 Flp-In T-REx cells and in anti tag coimmunoprecipitation; date restricted to January 2009 – January 2010.

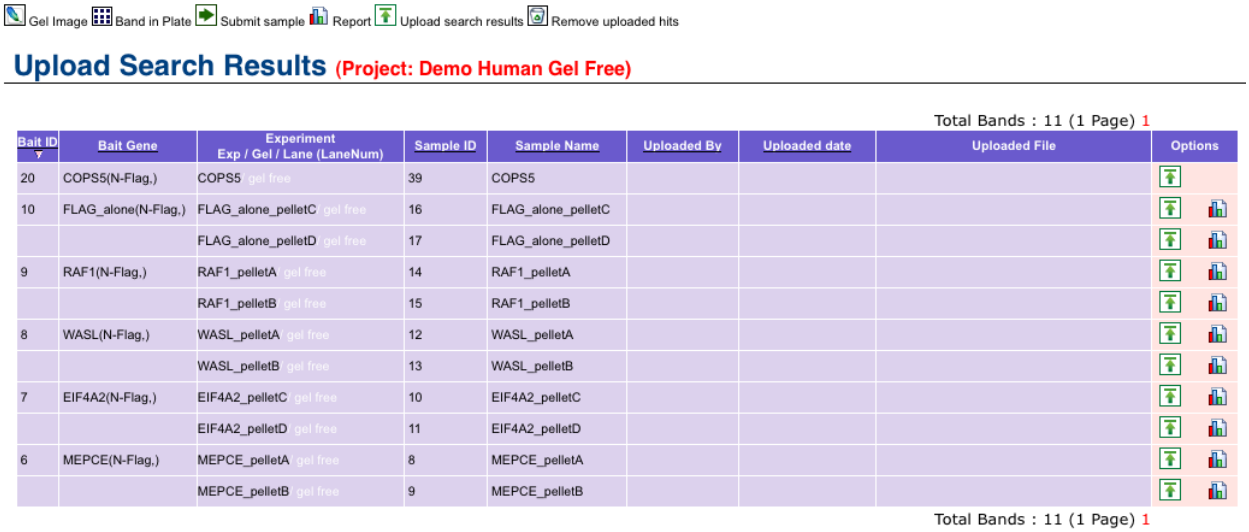


Uploading search results

The Analyst module allows you to import search results from the TransProteomics Pipeline (TPP), Mascot or GPM/X!Tandem. This function is very useful for laboratories that are not interested in the Data Management module of ProHits, e.g. if they are using a third party analysis solution. All that is needed for this section are the search results files or both TPP ProteinProphet and TPP PeptideProphet XML files.

⇒ Select the [Upload Search Results] link on the left hand side of the Analyst module.

This opens up the list of all the baits that you have created in the Analyst module.


Gel Image Band in Plate Submit sample Report Upload search results Remove uploaded hits

Upload Search Results (Project: Demo Human Gel Free)

Total Bands : 11 (1 Page) 1

Bait ID	Bait Gene	Experiment Exp / Gel / Lane (LaneNum)	Sample ID	Sample Name	Uploaded By	Uploaded date	Uploaded File	Options
20	COPS5(N-Flag,)	COPS5 gel free	39	COPS5				
10	FLAG_alone(N-Flag,)	FLAG_alone_pelletC gel free	16	FLAG_alone_pelletC				
		FLAG_alone_pelletD gel free	17	FLAG_alone_pelletD				
9	RAF1(N-Flag,)	RAF1_pelletA gel free	14	RAF1_pelletA				
		RAF1_pelletB gel free	15	RAF1_pelletB				
8	WASL(N-Flag,)	WASL_pelletA gel free	12	WASL_pelletA				
		WASL_pelletB gel free	13	WASL_pelletB				
7	EIF4A2(N-Flag,)	EIF4A2_pelletC gel free	10	EIF4A2_pelletC				
		EIF4A2_pelletD gel free	11	EIF4A2_pelletD				
6	MEPCE(N-Flag,)	MEPCE_pelletA gel free	8	MEPCE_pelletA				
		MEPCE_pelletB gel free	9	MEPCE_pelletB				

Total Bands : 11 (1 Page) 1

⇒ Select the upload option **at the end of the desired sample**

This pops up a new page:

Upload Search Results

Bait Information (20) [GI][Gene][BioGrid]

Bait Gene ID	10987	**Bait Gene Name**	COPS5 (N-Flag)
Bait Locus Tag		**Bait MW (kDa)**	37.560
Bait Clone	N/A	**Bait Description**	Jun activation domain binding protein [Homo sapiens]

Experiment Information

Experiment ID	Name/Batch Name	Exp. Detail	Exp. Status:	Inputed by	Date
32	COPS5	Interaction detection method : MI:0007 anti tag coimmunoprecipitation; Cell type : 293 Flp-In T-REx Invitrogen		Prohits Administrator	2010-01-15

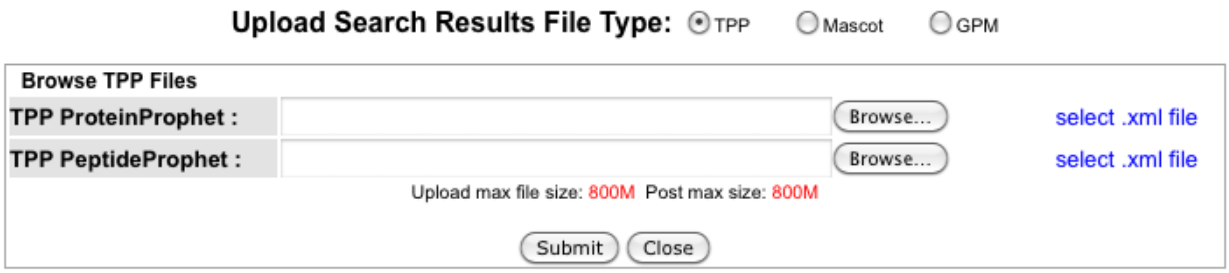
Sample Information

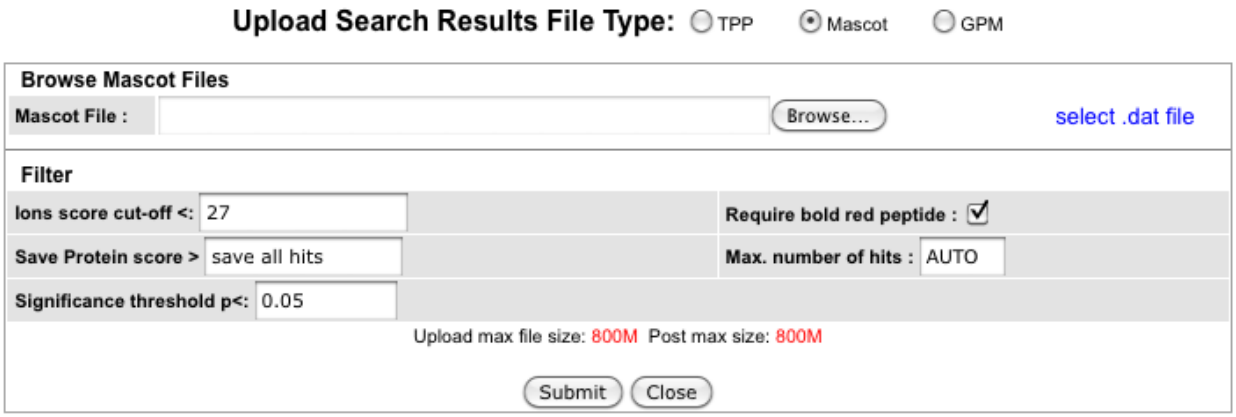
Sample ID	Sample Code	Submitted by
39	COPS5	Prohits Administrator

Upload Search Results File Type: ○ TPP ○ Mascot ○ GPM

⇒ Select the type of search results files you wish to upload (TPP, Mascot, GPM/X!Tandem), and Browse your local computer for the files in the right format.

⇒ Press [Submit] to upload search results.





Manage Protocols and Lists

Five types of Protocols and Lists pages are available in ProHits. With the exception of the "Epitope Tag Lists", that are applied to all projects on the local ProHits server, the other protocols and lists are only applicable to the current project. Lists and Protocols defined for a given project may be imported into a different project, so long as the user has access to both projects, and permission to modify individual lists and/or protocols. Access to individual pages of the "Manage Protocols and Lists" of the Analyst module is restricted via page permissions set in the admin office module. We suggest limiting the number of users having access to these management tools.

Here, we will briefly review the function of the different protocols and lists, then show a few examples for each category of protocol and/or list.

Text-based protocols – pages 42-43

Text-based protocols provide details on the experimental procedures. We have separated the protocols into four modules: Biological Material (i.e. what type of cells, expression system, growth conditions, etc.), Affinity Purification (from cell lysis to elution), Peptide Preparation (including separation at the protein/peptide level after elution), and LC-MS conditions. For our internal use, we attempt in providing very detailed protocols that could be used for publication with only minor modifications.

Experimental Editor – pages 44-45

The Experimental Editor allows you to create and manage the list of controlled vocabularies to be used within the Experimental Details page, in conjunction with the text-based protocols and additional notes. For our internal use, we attempt to capture information that would allow PSI MI 2.5 compliance, as well as other relevant information that would allow us to structure our data. Note that the terms entered in this section are searchable in the "Advanced Search" function.

Background Lists – pages 46-48

This function allows you to define and manage one or more lists of contaminants and/or background proteins associated with a given project. For example, you could maintain individual lists of the proteins found to associate non-specifically with different affinity matrices. The proteins on a given "Background" list can be subtracted from the list of identified proteins, in Individual Report, Comparison, or Export views.

Group Lists – pages 49-51

This function allows you to further organize and/or mark certain baits, experiments or samples by adding a colour-coded and user-defined icon that will appear in the Status bar of the Report by Bait or Report by Sample view. Useful Sample level group could include comments about the quality of the data, while Experiment level group would refer to some property of the experimental prep (e.g. phospho-enrichment), and a Bait level group could be the type of tag used. Additionally, ProHits allows you to mark (at the Sample level), samples that are to be included in publication (and/or to be exported to a third party).

Epitope Tag Lists – page 52

This is the only list that applies to the entire local ProHits database. The objects in this list are available on the Bait entry page, and define the tag (if applicable) used for tagging of the bait. N or C refer to the position of the tag relative to the bait. When available, the epitope tags have been mapped back to the standard vocabularies from the Molecular Interaction PSI MI 2.5; an automated link to the Ontology Lookup Service (OLS). We strongly suggest using this service to enter the PSI MI 2.5 terms when entering new tags.

Text-based protocols

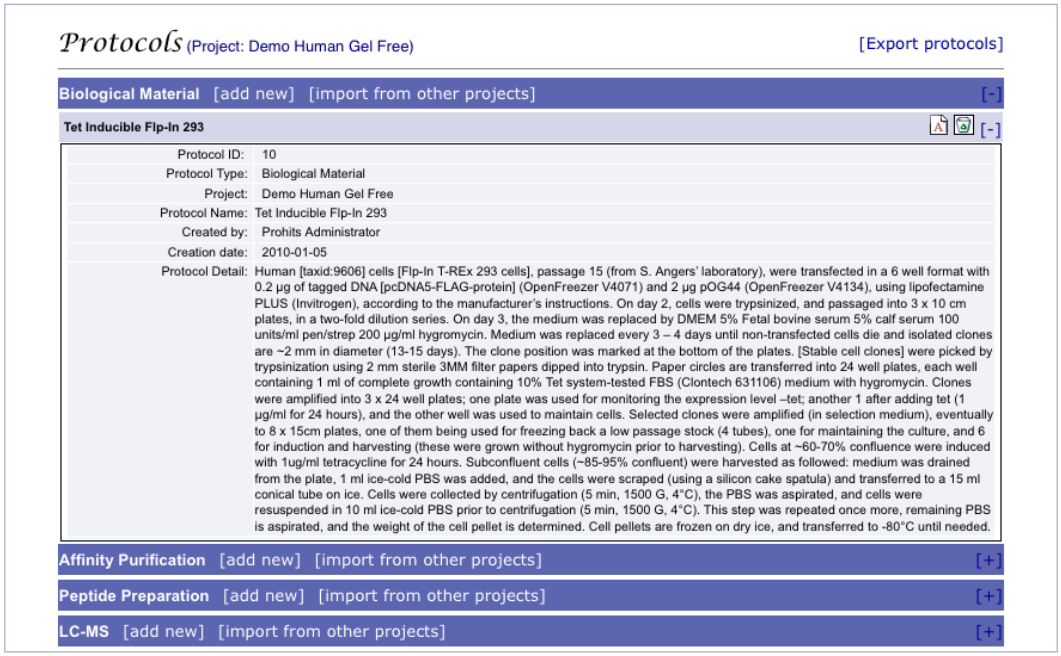
⇒ Select the "Text-based Protocols" entry from the Manage Protocols and Lists

Protocols (Project: Demo Human Gel Free)
[Export protocols]
Biological Material [add new] [import from other projects] [+]
Affinity Purification [add new] [import from other projects] [+]
Peptide Preparation [add new] [import from other projects] [+]
LC-MS [add new] [import from other projects] [+]

⇒ Click [add new], and paste or type your protocol. Then press [Save].

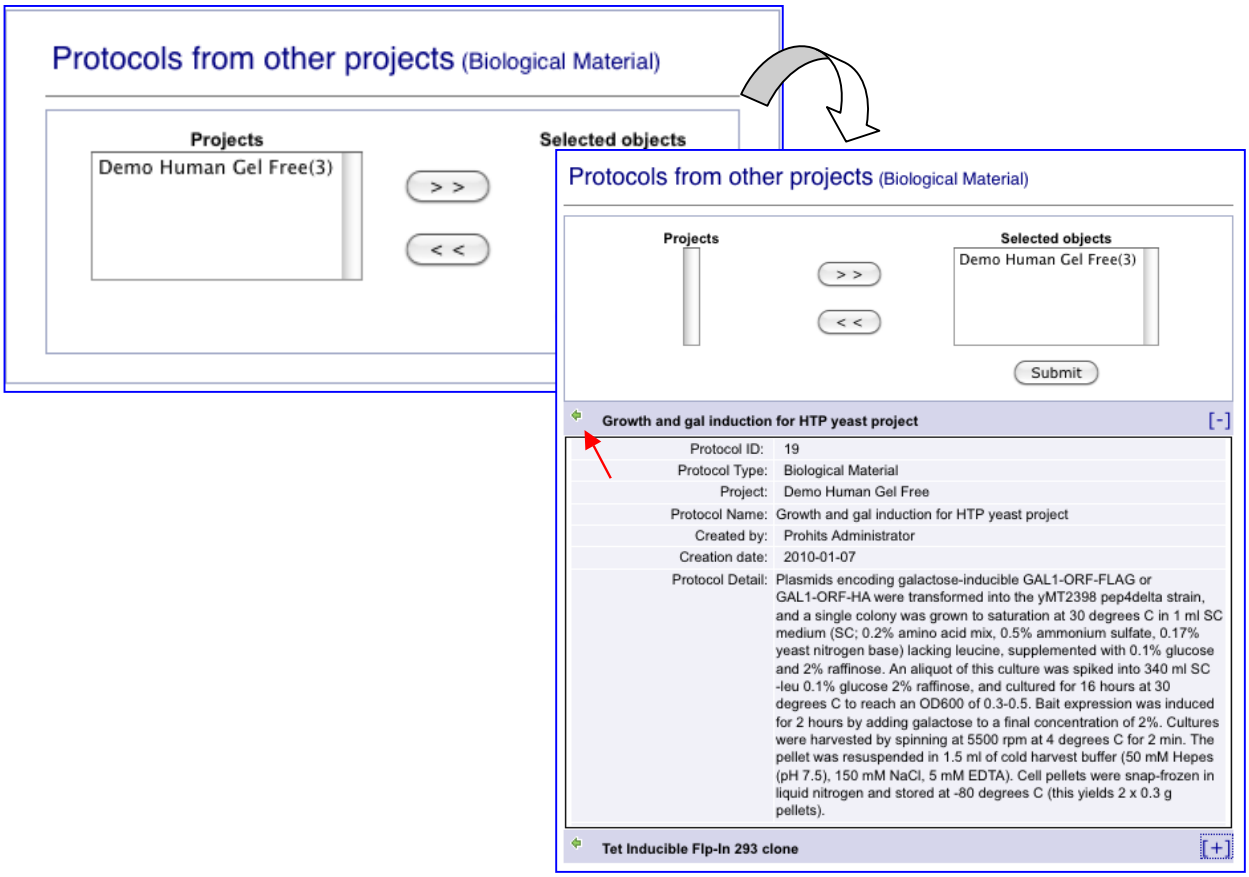
Note 1: because the protocols are displayed as html and exported as a CSV or TSV file, certain characters and symbols will not display properly, and should be spelled out. Examples are µ (u or micro), ° (degree), and ` (apostrophe).

Note 2: to each protocol is assigned a unique identifier. The protocol can be modified or even deleted as long as it has not been used. Once in use, modifications are no longer allowed and a new protocol (that will be assigned a different protocol number) will need to be created.


Protocols (Project: Demo Human Gel Free)
[Export protocols]
Biological Material [add new] [import from other projects] [-]
Tet Inducible Flp-In 293 [-]
Protocol ID: 10
Protocol Type: Biological Material
Project: Demo Human Gel Free
Protocol Name: Tet Inducible Flp-In 293
Created by: Prohits Administrator
Creation date: 2010-01-05
Protocol Detail: Human [taxid:9606] cells [Flp-In T-REx 293 cells], passage 15 (from S. Angers' laboratory), were transfected in a 6 well format with 0.2 µg of tagged DNA [pcDNA5-FLAG-protein] (OpenFreezer V4071) and 2 µg pOG44 (OpenFreezer V4134), using lipofectamine PLUS (Invitrogen), according to the manufacturer's instructions. On day 2, cells were trypsinized, and passaged into 3 x 10 cm plates, in a two-fold dilution series. On day 3, the medium was replaced by DMEM 5% Fetal bovine serum 5% calf serum 100 units/ml pen/strep 200 µg/ml hygromycin. Medium was replaced every 3 – 4 days until non-transfected cells die and isolated clones are ~2 mm in diameter (13-15 days). The clone position was marked at the bottom of the plates. [Stable cell clones] were picked by trypsinization using 2 mm sterile 3MM filter papers dipped into trypsin. Paper circles are transferred into 24 well plates, each well containing 1 ml of complete growth containing 10% Tet system-tested FBS (Clontech 631106) medium with hygromycin. Clones were amplified into 3 x 24 well plates; one plate was used for monitoring the expression level –tet; another 1 after adding tet (1 µg/ml for 24 hours), and the other well was used to maintain cells. Selected clones were amplified (in selection medium), eventually to 8 x 15cm plates, one of them being used for freezing back a low passage stock (4 tubes), one for maintaining the culture, and 6 for induction and harvesting (these were grown without hygromycin prior to harvesting). Cells at ~60-70% confluence were induced with 1ug/ml tetracycline for 24 hours. Subconfluent cells (~85-95% confluent) were harvested as followed: medium was drained from the plate, 1 ml ice-cold PBS was added, and the cells were scraped (using a silicon cake spatula) and transferred to a 15 ml conical tube on ice. Cells were collected by centrifugation (5 min, 1500 G, 4°C), the PBS was aspirated, and cells were resuspended in 10 ml ice-cold PBS prior to centrifugation (5 min, 1500 G, 4°C). This step was repeated once more, remaining PBS is aspirated, and the weight of the cell pellet is determined. Cell pellets are frozen on dry ice, and transferred to -80°C until needed.
Affinity Purification [add new] [import from other projects] [+]
Peptide Preparation [add new] [import from other projects] [+]
LC-MS [add new] [import from other projects] [+]

Continue entering protocols as above. Alternatively, if a protocol of interest already exists in another project to which you have access, you can import it directly from that project.

⇒ Click [import from other projects], select desired project by clicking the >> button and pressing [Submit].


Protocols from other projects (Biological Material)
Projects
Demo Human Gel Free(3)
> >
< <
Selected objects
Protocols from other projects (Biological Material)
Projects
> >
< <
Selected objects
Demo Human Gel Free(3)
Submit
Growth and gal induction for HTP yeast project
[-]
Protocol ID: 19
Protocol Type: Biological Material
Project: Demo Human Gel Free
Protocol Name: Growth and gal induction for HTP yeast project
Created by: Prohits Administrator
Creation date: 2010-01-07
Protocol Detail: Plasmids encoding galactose-inducible GAL1-ORF-FLAG or GAL1-ORF-HA were transformed into the yMT2398 pep4delta strain, and a single colony was grown to saturation at 30 degrees C in 1 ml SC medium (SC; 0.2% amino acid mix, 0.5% ammonium sulfate, 0.17% yeast nitrogen base) lacking leucine, supplemented with 0.1% glucose and 2% raffinose. An aliquot of this culture was spiked into 340 ml SC -leu 0.1% glucose 2% raffinose, and cultured for 16 hours at 30 degrees C to reach an OD600 of 0.3-0.5. Bait expression was induced for 2 hours by adding galactose to a final concentration of 2%. Cultures were harvested by spinning at 5500 rpm at 4 degrees C for 2 min. The pellet was resuspended in 1.5 ml of cold harvest buffer (50 mM Hepes (pH 7.5), 150 mM NaCl, 5 mM EDTA). Cell pellets were snap-frozen in liquid nitrogen and stored at -80 degrees C (this yields 2 x 0.3 g pellets).
Tet Inducible Flp-In 293 clone
[+]

⇒ Click the green arrow to transfer the protocol from the source project to the destination project, modify if needed, and press [Save].

You can export protocols linked to a project to a CSV file that can be opened in Excel or similar programs. The "Detail" column contains the full text of the protocol.

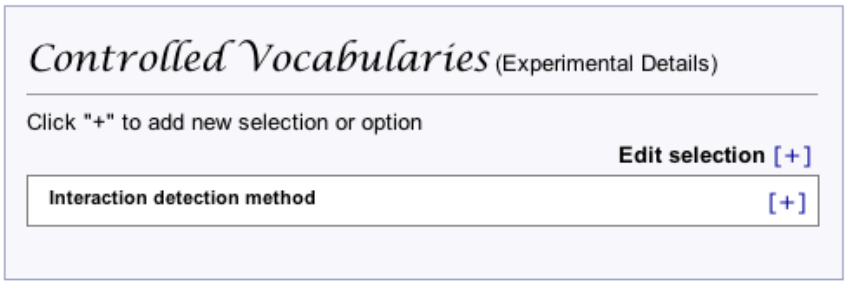
Sheets Charts SmartA

	A	B	C	D	E	F	G	H
	ID	Name	Type	Project Name	Detail	Creation Da	Creater	
	21	Trypsin digestion on magnetic beads	Peptide Preparation	Demo Human Gel Free	Proteins immobilize	1/7/10	Prohits Administrator	
	17	In-solution digest of IP samples	Peptide Preparation	Demo Human Gel Free	Trypsin (1 ug Sigma	1/7/10	Prohits Administrator	
	19	Growth and gal induction for HTP yeast project	Biological Material	Demo Human Gel Free	Plasmids encoding g	1/7/10	Prohits Administrator	
	15	Tet Inducible Flp-In 293 clone	Biological Material	Demo Human Gel Free	Human [taxid:9606	1/7/10	Prohits Administrator	

Experimental Editor

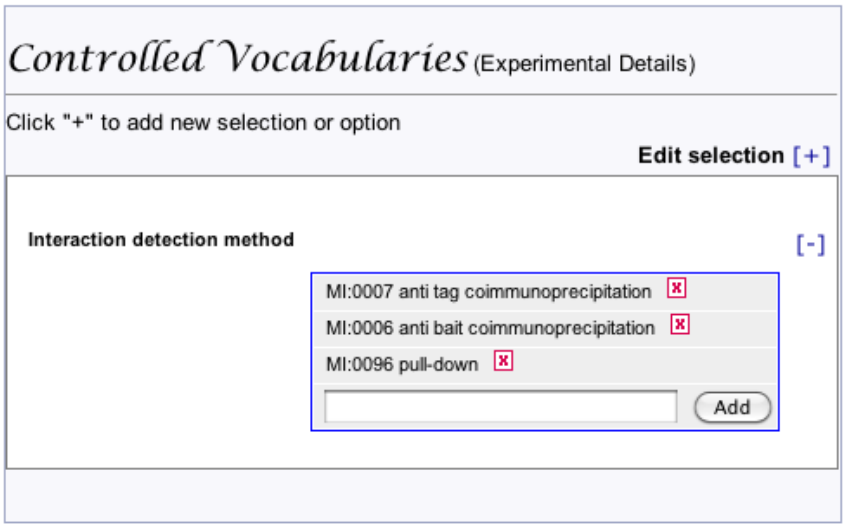
⇒ Select the "Experimental Editor" entry from the Manage Protocols and Lists

You will see a list of the categories already defined for your project.



⇒ To view the values already entered under the "interaction detection method" category, click on the [+] button to expand this category.

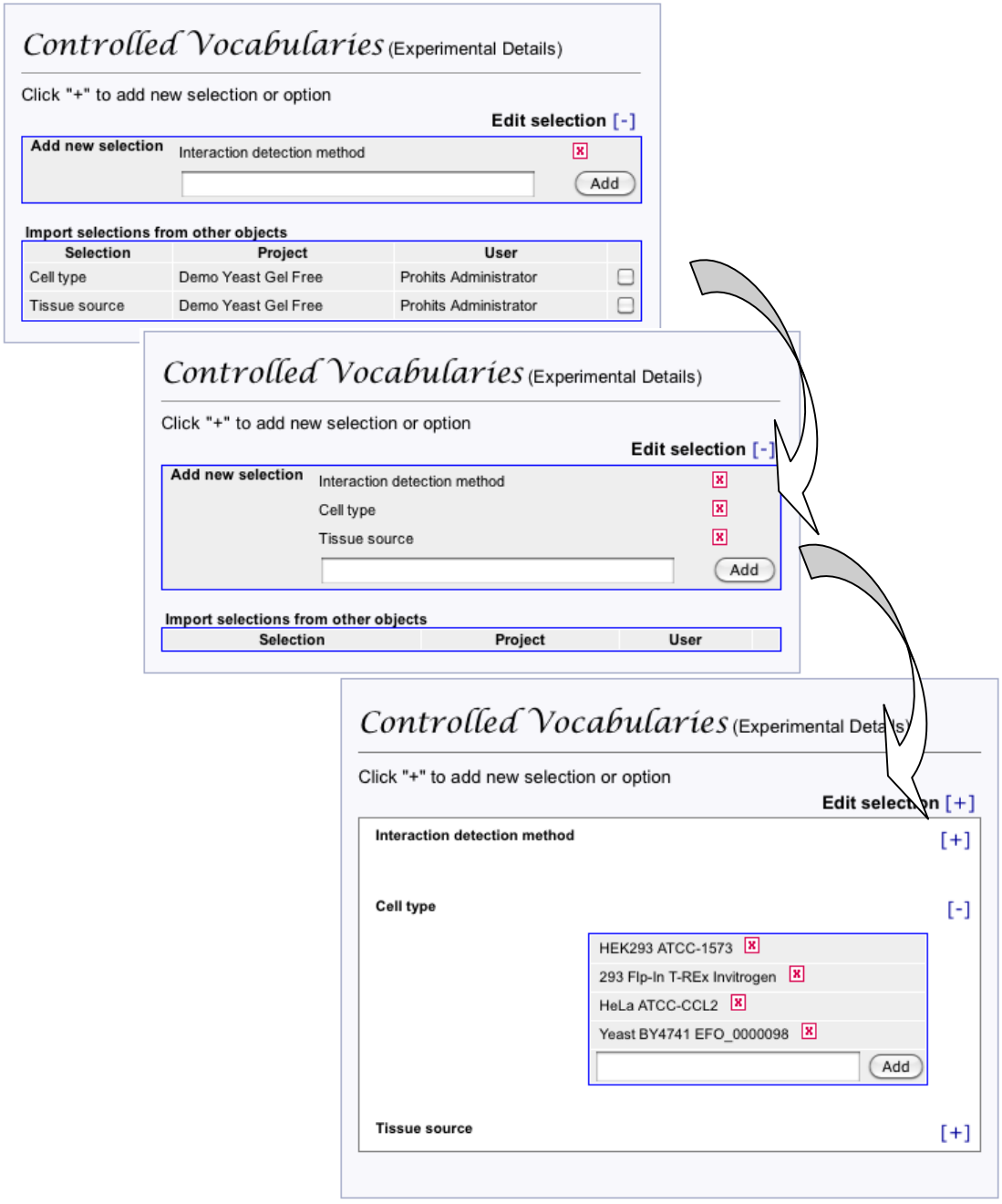
You can add additional values by typing their description and pressing [Add]. Values that are not yet linked to an entry are followed by a red **X**. Pressing **X** deletes the entry. Note that for this category, we have used PSI MI 2.5 terms, to facilitate later deposition in interaction databases.



⇒ To define new categories, press the [+] button next to "Edit selection".

This allows you to enter a new category.

⇒ To import a category from another project to which you have access, simply click the checkbox associated to the category under the Edit selection option to transfer the category (and associated values) to current project.

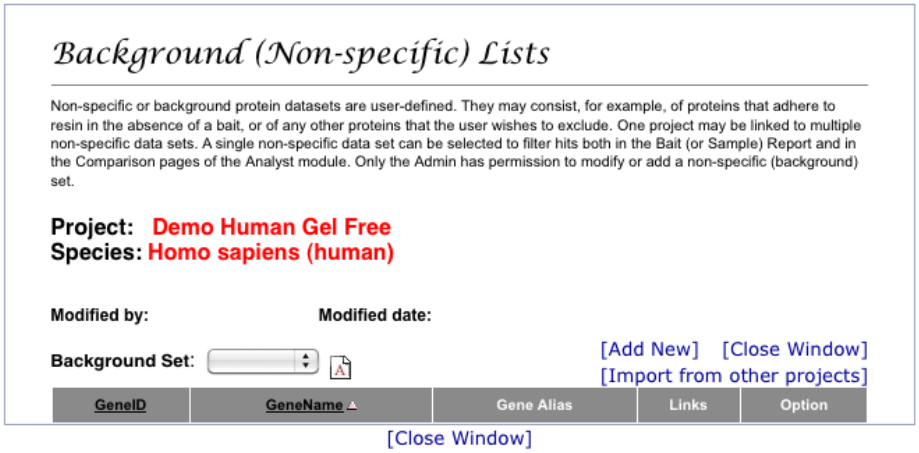


Background Lists

In addition to the Bio Filters and Experimental Filters defined in the Admin Office module, ProHits allows you to define additional filters to remove non-specific (or background) proteins. These filters are project-specific and created within a bait (or sample) report page in the Analyst module. Several different filters can be associated with the same project (e.g. corresponding to different workflows used in the project). Creation of these filters requires administrator-level privileges. The filters can be created by adding proteins manually (one-by-one) to an existing list of contaminants. The filters can also be generated by uploading a list (or table) of hits identified in control run(s), in which case the mapping only requires the Entrez Gene ID field. You can also add multiple proteins at once from any other pre-existing list (e.g. in Excel). The mapping is via the NCBI Entrez Gene ID.

⇒ Select the "Background Lists" entry from the Manage Protocols and Lists.

⇒ From the entry page, click on the (modify) icon to upload a list of contaminant proteins.


Background (Non-specific) Lists
Non-specific or background protein datasets are user-defined. They may consist, for example, of proteins that adhere to resin in the absence of a bait, or of any other proteins that the user wishes to exclude. One project may be linked to multiple non-specific data sets. A single non-specific data set can be selected to filter hits both in the Bait (or Sample) Report and in the Comparison pages of the Analyst module. Only the Admin has permission to modify or add a non-specific (background) set.
Project: Demo Human Gel Free
Species: Homo sapiens (human)
Modified by:
Modified date:
Background Set:
[Add New] [Close Window]
[Import from other projects]
GeneID
GeneName
Gene Alias
Links
Option
[Close Window]

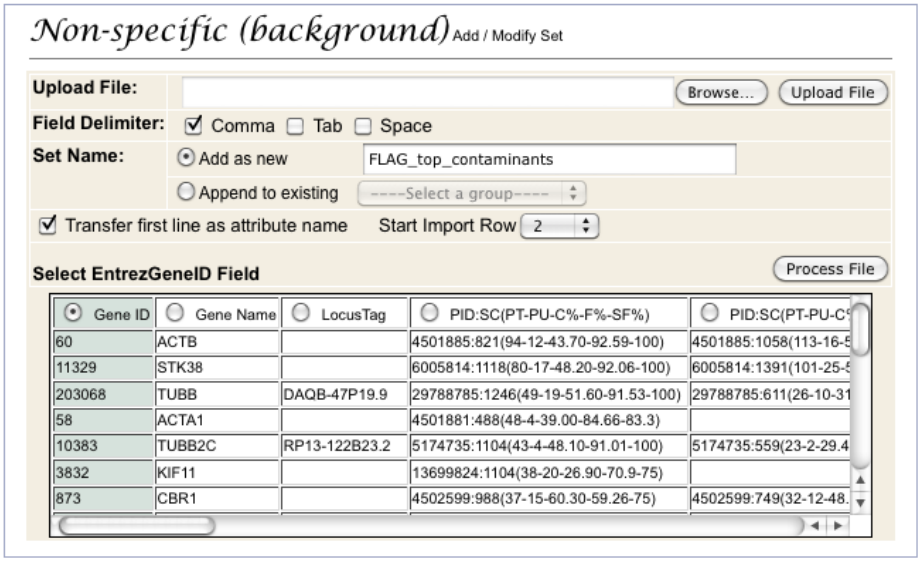
Note that an efficient method to generate a non-specific filter set utilizes the ProHits comparison tool. First, select multiple control runs and merge them into a single "Control" group. This will open up a Comparison page with a single column called "Control Group" displayed in yellow. As before, the maximal value for the parameter visualized is displayed (e.g. spectral count). Apply filters (e.g. number of unique peptides, protein coverage, etc.) desired, and select [Export(table)] to export a comma-delimited file (*.csv). Save this file on your hard drive, and go to any Bait report page. (Note that any Excel or text file that lists the NCBI Gene ID may also be used).

⇒ Browse the file to be uploaded, select delimiter, and press [upload file].

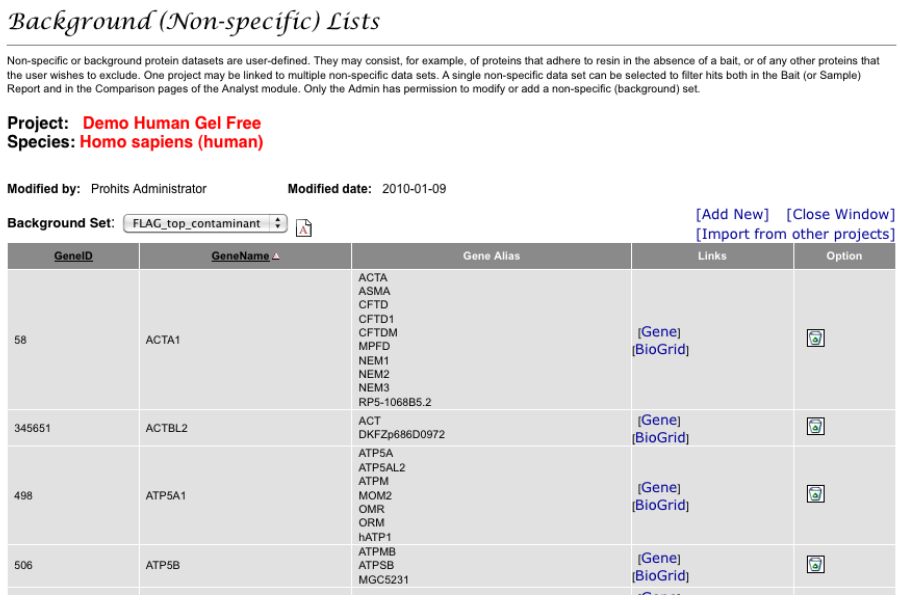
⇒ Select the "add as new" radio button and type a name (here: FLAG_top_contaminants).

Alternatively, append to an existing list by using the dropdown menu.

⇒ Select the row to start importing, and check the radio button in the GeneID field. Then, click [Process File].

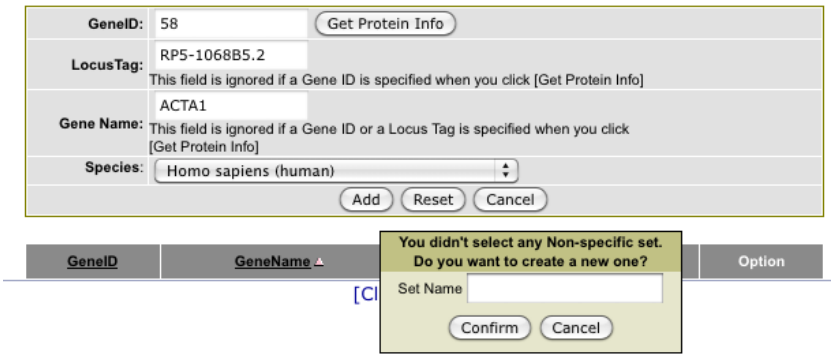


Once the file is processed, the contaminant list will be displayed (after selecting the name in the dropdown menu). You can manually remove individual entries (they will not be on the background list) by clicking the "delete" icon.


Background (Non-specific) Lists
Non-specific or background protein datasets are user-defined. They may consist, for example, of proteins that adhere to resin in the absence of a bait, or of any other proteins that the user wishes to exclude. One project may be linked to multiple non-specific data sets. A single non-specific data set can be selected to filter hits both in the Bait (or Sample) Report and in the Comparison pages of the Analyst module. Only the Admin has permission to modify or add a non-specific (background) set.
Project: Demo Human Gel Free
Species: Homo sapiens (human)
Modified by: Prohits Administrator
Modified date: 2010-01-09
Background Set: FLAG_top_contaminant
[Add New] [Close Window]
[Import from other projects]
GeneID GeneName Gene Alias Links Option
58 ACTA1 ACTA ASMA CFTD CFTD1 CFTDM MPFD NEM1 NEM2 NEM3 RP5-1068B5.2 [Gene] [BioGrid]
345651 ACTBL2 ACT DKFZp686D0972 [Gene] [BioGrid]
498 ATP5A1 ATP5A ATP5AL2 ATPM MOM2 OMR ORM hATP1 [Gene] [BioGrid]
506 ATP5B ATPMB ATPSB MGC5231 [Gene] [BioGrid]

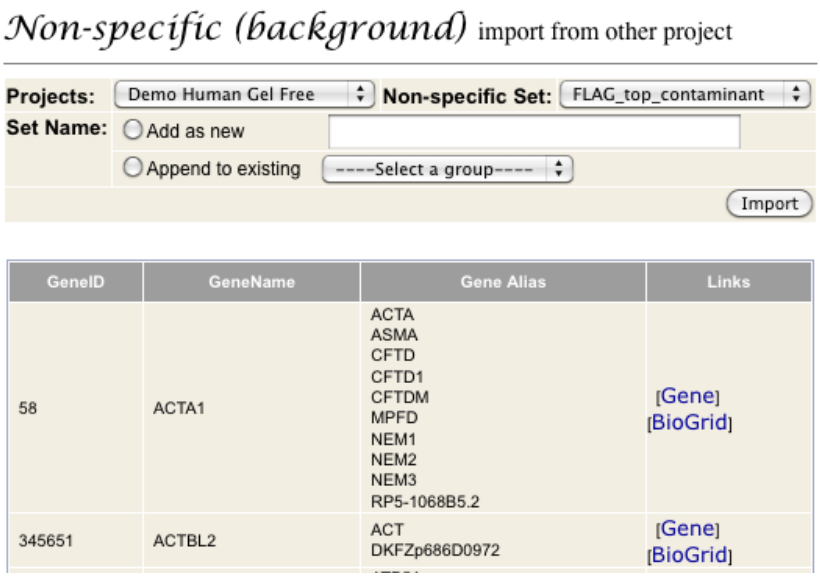
⇒ To manually add a protein to a background list, press [Add New].

You will then be prompted to enter a new contaminating/background protein. You can simply enter a gene name and species and press [Get Protein Info]. Press [Add] to include this protein on the background list.



If you do not specify a pre-entered non-specific set, ProHits will allow you to create a new one (press [Confirm] after entering the non-specific set name).

⇒ To import a contaminant list from a different project, press [Import from other projects], and navigate through the menus.



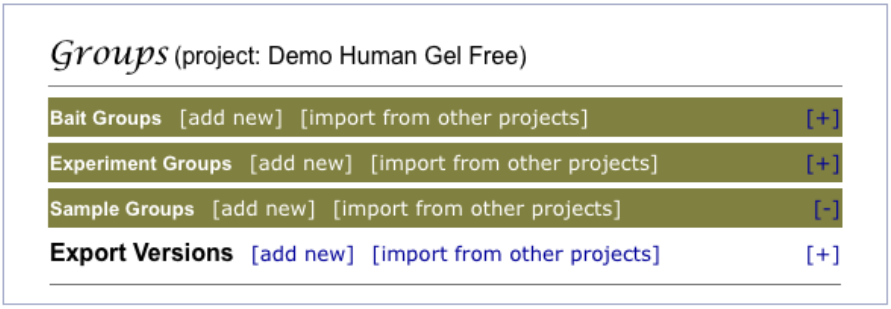
You now have your own background set that can be used for filtering both in the bait/sample report pages and in comparison. We recommend using caution when creating these sets: some proteins that are true interacting partners for a given bait may also be present (usually in lower amounts) on the background list. It may be a good idea to only include on this non-specific (background) list proteins detected across more than one control run with a high number of peptides.

Group Lists

ProHits allows the definition of new "groups" for any given project. As described earlier, groups are added to baits/samples by selecting the "Notes" Option. Groups act like flags and are displayed in the status bar in the "Report by Bait" or "Report by Sample" pages. These groups can help you organize your data.

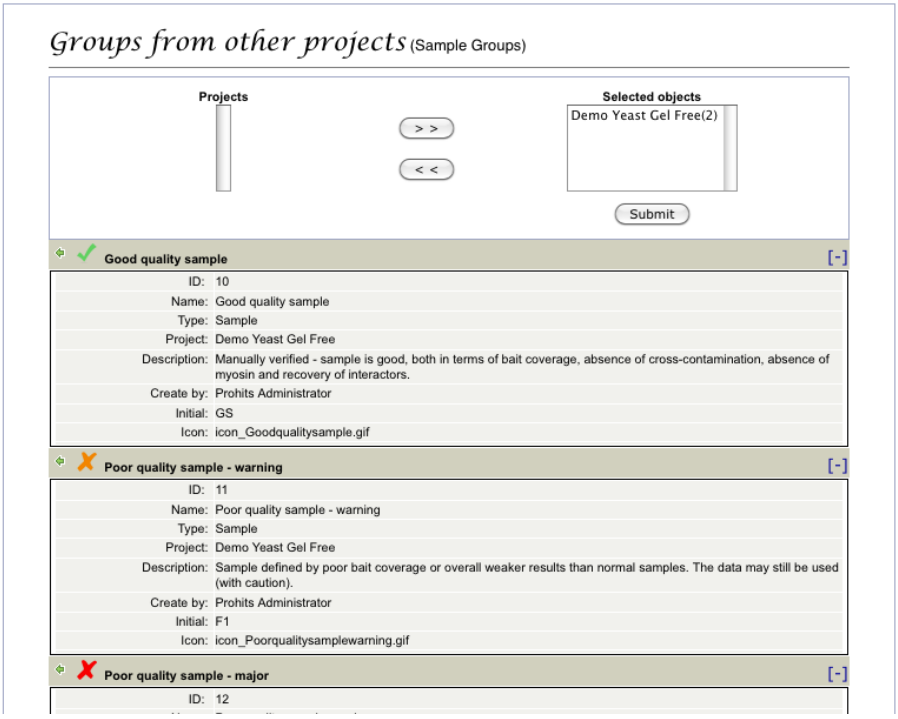
⇒ Select the "Groups" entry from the Manage Protocols and Lists.

As with the other Protocols and Lists, you can define new groups, or import a new group from another project. Here we will import sample groups from a different project.

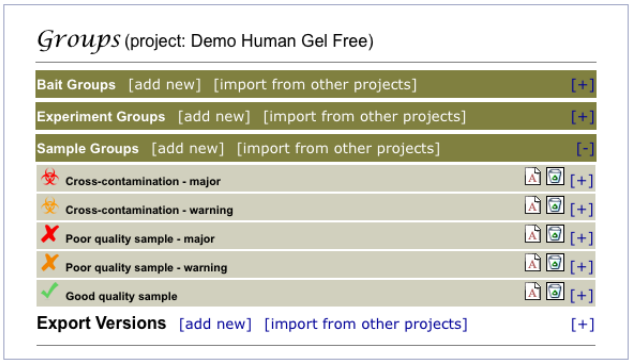

Groups (project: Demo Human Gel Free)
Bait Groups [add new] [import from other projects] [+]
Experiment Groups [add new] [import from other projects] [+]
Sample Groups [add new] [import from other projects] [-]
Export Versions [add new] [import from other projects] [+]

⇒ To import groups from a different project, press [import from other projects], and navigate through the menus.

As with the Text-based protocols, use the green arrows to transfer desired groups to the current project. You can only transfer one group at a time.

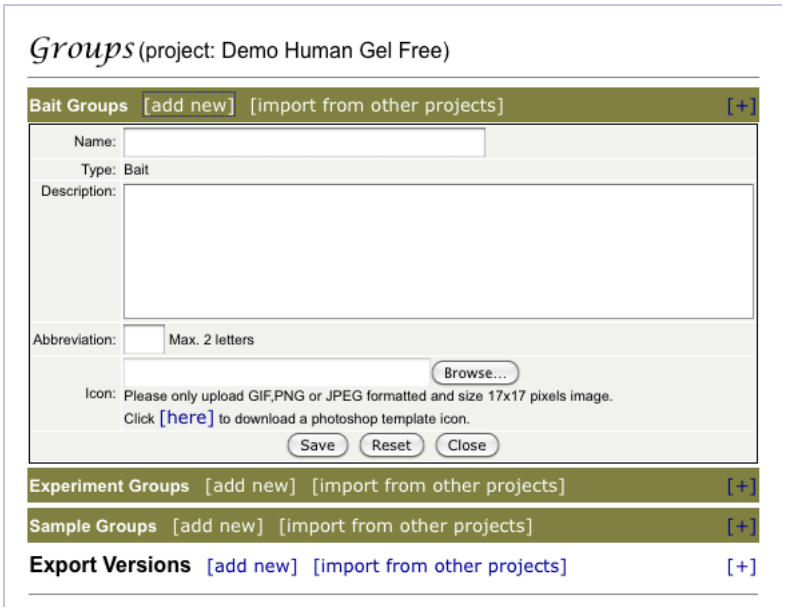

Groups from other projects (Sample Groups)
Projects
> >
< <
Selected objects
Demo Yeast Gel Free(2)
Submit
Good quality sample [-]
ID: 10
Name: Good quality sample
Type: Sample
Project: Demo Yeast Gel Free
Description: Manually verified - sample is good, both in terms of bait coverage, absence of cross-contamination, absence of myosin and recovery of interactors.
Create by: Prohits Administrator
Initial: GS
Icon: icon_Goodqualitysample.gif
Poor quality sample - warning [-]
ID: 11
Name: Poor quality sample - warning
Type: Sample
Project: Demo Yeast Gel Free
Description: Sample defined by poor bait coverage or overall weaker results than normal samples. The data may still be used (with caution).
Create by: Prohits Administrator
Initial: F1
Icon: icon_Poorqualitysamplewarning.gif
Poor quality sample - major [-]
ID: 12

Upon transfer of a group, it will appear on your group list as shown below (the new group can be modified or deleted, unless it is used for a sample).


Groups (project: Demo Human Gel Free)
Bait Groups [add new] [import from other projects] [+]
Experiment Groups [add new] [import from other projects] [+]
Sample Groups [add new] [import from other projects] [-]
Cross-contamination - major [+]
Cross-contamination - warning [+]
Poor quality sample - major [+]
Poor quality sample - warning [+]
Good quality sample [+]
Export Versions [add new] [import from other projects] [+]

⇒ To create a new group, press [add new], and navigate through the menus.

Simply enter a short descriptive name for the group as well as a description, an abbreviation (that will be listed alongside the baits or samples), and an icon. Icons can easily be created in Photoshop as 17 x 17 pixel images, and saved as GIF, PNG or JPEG files. A template can be downloaded from the ProHits group page.


Groups (project: Demo Human Gel Free)
Bait Groups [add new] [import from other projects] [+]
Name:
Type: Bait
Description:
Abbreviation: Max. 2 letters
Browse...
Icon: Please only upload GIF,PNG or JPEG formatted and size 17x17 pixels image.
Click [here] to download a photoshop template icon.
Save Reset Close
Experiment Groups [add new] [import from other projects] [+]
Sample Groups [add new] [import from other projects] [+]
Export Versions [add new] [import from other projects] [+]

Export version

ProHits allows you to flag a group of samples, e.g. for inclusion in a publication or export to a third party.

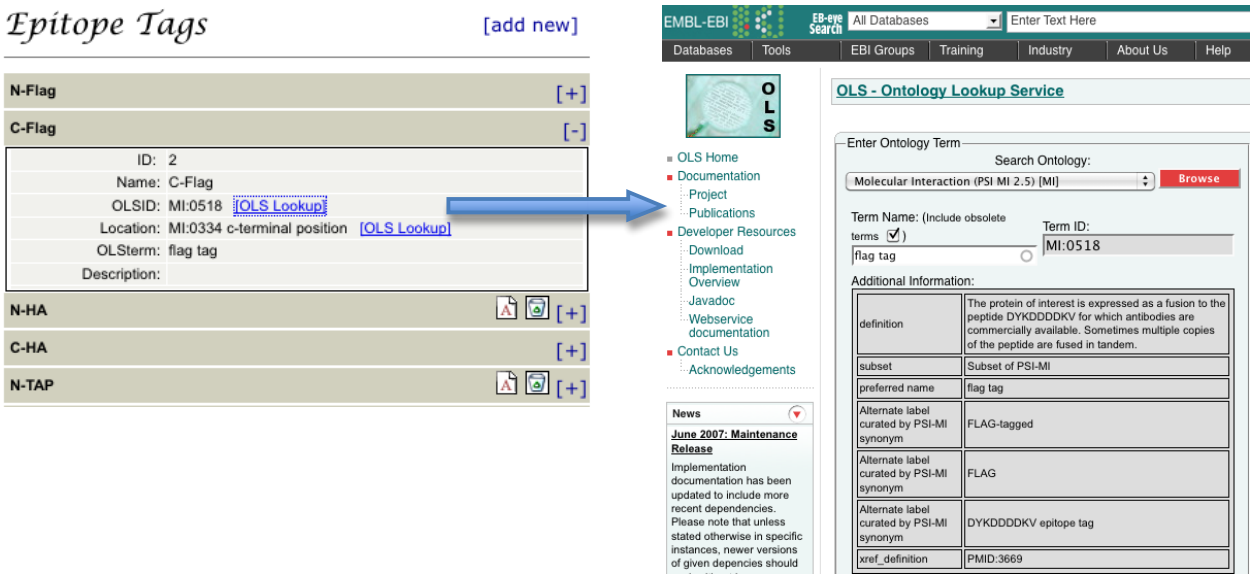
⇒ To create an Export Version, press [add new].

This will open a new menu with the default abbreviation (Version1, VS1), and Icon (a yellow star with the number 1). Subsequent versions will automatically be numbered VS2, VS3, etc., and the number inside the star will similarly increase. We suggest that you provide a meaningful short name and an accurate description of each "Export Version".

Epitope Tag Lists

⇒ Select the "Epitope Tag Lists" entry from the Manage Protocols and Lists.

This lists all tags available to the local ProHits projects. Clicking on the [+] sign expands the details of the epitope tag. We have mapped the current epitope tags in the demo database to PSI MI 2.5, using the Ontology Lookup Service (OLS) at the EBI. A link page is provided that allow retrieval of additional information.



In addition to the epitope tags currently in the system, you can create additional tags by pressing [add new] and navigating through the fields. Again, we strongly recommend mapping your terms to PSI MI 2.5 whenever possible.